2000

Cancer Mortality and *Environment in Louisiana: a Geographical Inquiry.

Mihye Bark
Louisiana State University and Agricultural & Mechanical College

CANCER MORTALITY AND ENVIRONMENT IN LOUISIANA: A GEOGRAPHICAL INQUIRY

A Dissertation

Submitted to the Graduate Faculty of the
Louisiana States University and
Agricultural and Mechanical College
in partial fulfillment of the
requirements for the degree of
Doctor of Philosophy

in

The Department of Geography and Anthropology

by
Mihye Bark
B.S., Kyungpook National University, 1986
M.S., Kyungpook National University, 1988
May 2000

ACKNOWLEDGEMENTS

First of all, I owe a tremendous debt of gratitude of my major Professor Nina S-N Lam, who devoted her time and expertise to guide me through what seemed the longest journey of my life. There are many people who had positive and valuable influences on me during my experience at Louisiana State University. I am especially indebted to the following committee members: Professors Margaret A. Reams, Kam-Biu Liu, William V. Davidson, and Robert W. Adkinson.

I express my appreciation to Mr. Holmes and his wife, whose encouragement was positive catalysts in motivating my thesis. I say a special prayer for Mrs. Holmes, who recently passed away. I can not forget my mother's constant love and support. I also thank my father, Park Sang-Po, and parents-in-law, Nho Kun-Ha and Bae Boon-Sun, for their encouragement and support during this rather long journey. I would like to express my special thanks to my husband, Sung Chul No, and my son, Joonwoo for their patient support.

Finally, I attribute all my works in life to a strong faith in God, who taught me that all things are possible through Christ, who strengthens me.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

ABSTRACT

Cancer is one of the most serious health problems in the U.S. Louisiana mortality rates for all sites of cancer have exceeded national rates since 1950. Even though concern about the issues of environmental cancer is growing, conclusive results have not been reported. Previous studies differed in data source, data quality, time of observation, and statistical procedure, making comparison of results problematic.

The spatial patterns of the most common cancer sites in Louisiana (from 1953 to1993) and their relationships with environmental factors were investigated, using GIS software and statistical software. To provide a clear understanding of how the spatial distributions of cancer mortality rates in Louisiana compared with those in the nation, this research was conducted using data at different geographic levels.

Cancer mortality rates for most parishes in South Louisiana were higher than those of the entire U.S. for all sites combined as well as lung and stomach cancer, individually. Spatial clusterings of cancer mortality rates were tested using factor analysis, spatial autocorrelation and correlograms, and scan statistic. Cancers of all sites combined, lung, and stomach exhibited the strongest positive spatial autocorrelation. In searching for the presence of geographical clusters for lung cancer deaths at parish and census tract levels from 1988 to 1993, results revealed that there was a statistically significant and geographically distinct cluster of lung cancer deaths in southeastern Louisiana from 1988 to 1993.

Stepwise regression analyses were applied to data to examine the relationships between environmental factors and cancer mortality rates. Results showed a positive relationship between cancer (all sites combined) and urban population and a negative

relationship between cancer and persons employed in health and education services. Breast cancer mortality rates were significantly positively related to urban residence and education level. Colorectal cancer mortality rates were associated with the type of drinking water. Lung cancer mortality rates were more closely related to occupational variables and agricultural chemicals than any other types of cancer mortality rates. Prostate and stomach cancer mortality rates were positively associated with nonwhite populations.

CHAPTER 1

INTRODUCTION

Human health research is concerned with how human activity and health condition interact with their physical, social, and cultural environments. Hippocrates commented upon the importance of this interaction between cultural-environmental conditions and human health more than 2,000 years ago (Mead et al. 1988). To understand human health, the interaction among human, environment, and disease must be considered.

Cancer is one of the leading causes of death in the United States. In 1990, the number of cancer deaths was 505,322, second (23.5%) to heart disease among all causes of death in the U.S. (Boring et al. 1993). Louisiana, particularly South Louisiana, has unusually high rates of certain types of cancer, so cancer is one of the most serious health problems in the state (Task Force on Environmental Health 1984). It is generally accepted that approximately 80% of all cancers have an environmental component (Higginson 1993). Many researchers think that environmental factors cause most cancers and therefore, cancers are theoretically preventable, once the causal factors are identified (Wynder et al. 1977; Doll et al. 1981; Higginson et al. 1992). It has been recognized that specific factors do not contribute independently but interdependently to the etiology in chronic diseases. It is natural to assume that there exist multiple factors which interact together in the pathogenesis of cancer of various sites.

Many studies have been undertaken to gain insight into the etiology of cancer and the spatial analysis between cancer and environment (Hoover and Fraumeni 1975; Glick 1977, 1979a, 1979b, 1982; Babin 1979; Inaba et al. 1981; Minowa et al. 1981; Lam 1986; Kennedy 1988; Openshaw et al. 1988; Shannon and Pyle 1992; Gatrell et al. 1995). Wong

and Foliart argued (1989) that studies exploring potential environmental cancer risks in Louisiana were very limited. These studies produced some theories about possible causes, but no solid answers.

The geographical distribution and variations of mortality, incidence, and prevalence of diseases have proven valuable in generating hypotheses of disease etiology. Furthermore, the importance of Geographic Information Systems (GIS) in the study of geography and related disciplines has increased dramatically over the last few years. There is much potential for the use of GIS in the management and analysis of health care, health, and environmental factors. Therefore, it is crucial that its use, in analyzing the geographical distributions of cancer and environmental factors, be examined.

This research reported here involves the relationship of geographic distribution of cancer with environmental factors for the most common cancer sites in Louisiana from the 1950s to the 1990s. Statistical mapping, significance test of rate difference, factor analysis, regression analysis, spatial autocorrelation analysis, and scan statistic are applied to examine the spatial patterns of cancer and their relationships with environmental factors. GIS software and statistical software are used to store and manage large amounts of cancer data and environmental parameters as well as to map, display, analyze, and visualize the spatial patterns of the environment-related cancer and its causes.

1.1 Problem Statement

Mortality rates for all cancer sites in Louisiana, particularly South Louisiana, are among the highest in the nation. The state ranked number 10 in the U.S. mortality rates from cancer at all sites from 1950 to 1959, number 8 from 1960 to 1969, number 4 from 1970 to 1979 (Riggan et al. 1983), number 4 from 1980 to 1989, and number in 3 from

1990 to 1994 (CDC Wonder website *http://wonder.cdc.gov/wonder/data/mortJ.shtml 1999*). Lung cancer mortality rates in white males have been among the highest in the nation (Riggan et al. 1983; CDC Wonder website *http://wonder.cdc.gov/wonder/data/mortJ.shtml* 1999).

Though concern about the issues of environmental cancer continues to grow and previous studies have shown some relationships between the distribution of cancer and environmental factors in Louisiana, definitive and conclusive results have not yet been reported. A few statistical or descriptive studies have indicated a relationship between cancer and environment, but conclusions and results have not been consistent. According to recent studies (Chen et al. 1990, 1991, 1996, 1997, 1998; Groves et al. 1996), cancer incidence rates in South Louisiana are generally similar to or lower than the Surveillance, Epidemiology, and End Results (SEER) rates.[1.1] On the other hand, mortality rates for cancers of all sites combined continue to exceed national rates. This discrepancy deserves more study and action.

The geographic distribution of disease has been analyzed in the hope of linking disease manifestation with specific environmental factors or with groups of factors. Spatial analytical techniques of disease patterns can be used as one tool to detect factors of disease causation. If the cause and nature of certain cancers attributed to environmental factors are known, preventative guidelines can be offered for many tumors.

[1.1] Surveillance, Epidemiology and End Results of the National Cancer Institute, a federal program which collects and analyzes information on cancer cases from 4 cities and 5 states of the country representing about 10% of the U.S. population.

1.2 Research Objectives

The goal of the study was to characterize the spatial patterns of the most common cancer types in Louisiana and possible relationships with environmental factors. Specifically, the objectives of the study were as follows:

1. To document the spatial distribution and temporal changes of national cancer mortality rates and to provide a clear understanding of how cancer mortality rates and trends in Louisiana compare with those in the nation.

2. To document the spatial distributions and temporal changes in cancer mortality rates in Louisiana.

3. To search for spatial patterns that might suggest an environmental etiology or causative hypotheses.

4. To identify the relationship between cancer and environmental factors in Louisiana.

5. To evaluate the use of GIS in studying cancer and associated environmental factors.

1.3 Research Hypotheses

Geographical concentrations of the disease could lead to hypotheses about local demographic or environmental factors in its causation (Dent and Goulston 1982). Cancer mortality maps and studies done in Louisiana have suggested hypotheses. However, these studies to date have not reached the degree of sophistication to allow testing of cause-effect hypotheses. Many of these hypotheses-generation studies have been interpreted only to imply correlation but not causation. Therefore, these hypotheses should be explored further.

The general hypotheses of the current research were as follows:

1. Cancer mortality rates (breast, colon and rectum, lung, prostate, and stomach) are higher in South Louisiana than in the nation or the state.

2. There are spatial clusterings of cancer mortality rates of some major sites including breast, colon and rectum, lung, prostate, and stomach.

3. These cancer mortality patterns can be associated with environmental variables, including Toxic Release Inventory (such as air, water, underground injection, and on-site land toxic releases), income, population, population density, occupation (such as agriculture, manufacturing of all products, chemical manufacturing, mining, and education), persons below poverty level, educational status (percent of high school graduates), smoking, alcohol consumption, number of solid and hazardous waste sites, distance from superfund sites, drinking water, and wetland areas.

Methods used were; statistical mapping, significance test of rate difference, and factor analysis for hypothesis 1; factor analysis, spatial autocorrelation analysis, and scan statistic analysis for hypothesis 2; and factor analysis and multiple regression analysis for hypothesis 3. Research methods are presented in detail in Chapter 3.

1.4 Expected Significance

The relationship of spatial distribution of cancer and environmental factors in Louisiana from the 1950s through the 1990s is investigated. The expected significance of this research extends to four major areas.

1. To identify possible causes of cancer attributed to environmental factors and to suggest preventative guidelines to monitor cancer risk factors.

2. To provide a systematic overview of the trends and patterns of cancer mortality rates in Louisiana.

3. To identify some promising methodological approaches to the study of the spatial distribution of cancer and environment.

4. To demonstrate the potential of using GIS to analyze medical phenomena, such as the ability to store, manage, map, integrate, analyze, and visualize a large cancer database and environmental elements.

5. To examine the strength of the evidence produced thus far in support of hypotheses about potential environmental cancer risks in Louisiana.

1.5 Overview of Research

Fundamental literature concerning the issues of cancer and environment is outlined in Chapter 2. The cancer mortality trends from international and national perspectives and the case in Louisiana are also discussed. Study area, data source, methods, and limitations of research are described in Chapter 3. Chapter 4 provides a base line for geographic patterns of cancer mortality rates in the U.S. during the years 1953-1987. Chapter 5 characterizes the geographic patterns of cancer mortality rates in Louisiana parishes from the 1950s to the 1980s, and examines the relationship between cancer mortality and environmental factors in the 1980s. Using the scan statistic, Chapter 6 examines and compares the geographic clusters of lung cancer death at different levels (parishes and census tracts) from 1988 to 1993. Finally, conclusions for this study and suggestions for further research are presented in Chapter 7.

CHAPTER 2

REVIEW OF LITERATURE

2.1 Human, Disease, and Environment

The important relationship of human health or disease with environmental factors was recognized long ago. While Hippocrates emphasized the role of environmental factors in human disease causation, little real data was available at the time (Higginson et al. 1992). Medical geography, which treats the provisions of human health, disease, health care, and environmental conditions, has been studied since the 18th century. It involves the description of spatial distributions of morbidity, mortality, and health care. Also, it considers possible causative relationships among sickness, disease, death, and local variations in environmental conditions (Goodall 1987).

In relationship to geography and health, the "environment" includes the biophysical, socioeconomic, and cultural worlds. Howe expressed this view:

> Health problems are environmental problems. The environment itself is a matrix of physical, biological, and socio-cultural circumstances which surround man and affect his physical, mental, and social well-being. It is the sum total of his habitat, economy and society and as such embraces not only his life support systems of air, water, food and shelter but also the multiplicity of factors which bear down on him and affect his general well-being (Howe 1986a, p. 388).

Cancer is a term used for a group of diseases characterized by uncontrolled growth and spread of abnormal cells, and applies to more than 100 diseases. There are multiple causes and treatments for cancer. Although cancer was known to the ancient Egyptians, nothing was known of the causes or geographical distribution of the disease. Knowledge and classification of cancer is analogous to our knowledge of infectious diseases in the early 19th century. When the Society for Investigating the Nature and Cure of Cancer

published findings about the disease in the Edinburgh Medical and Surgical Journal (1806), it commented:

> With regard to cancer, it is not only necessary to observe the effects of climate and local situation, but to extend our views of different employments, as those in various metals and manufactures; in mines and collieries; in the married or single; in different sex, and many other circumstances (Higginson et al. 1992, p. xviii).

The first significant attempt to determine the geographical distribution of cancer and its possible causes appeared in the 19th century. There have since been quite a few studies of the geographic distribution of cancer to provide clues to its causation or environmental factors. For example, Kennaway's (1944) study about liver cancer in the Negro in Africa and America found that blacks in southern Africa had a very high incidence of liver cancer, whereas blacks resident in the U.S. did not. It concluded that the disease was probably affected by some extrinsic factors in the environment. Howe (1960) was one of the first to map age-standardized mortality rates on the geographical distribution of cancer mortality in Wales. He identified possible environmental factors by visually comparing cancer maps to similar maps of air pollution, water supply, and other hypothesized cancer correlates. Murray (1962) mapped the distribution of mortality rates for some of the major causes of death, including cancer with external information such as the location of urban development, pollution sources, and soil types. Doll (1969) indicated that spatial variation in cancer rates suggests the importance of environmental influences in cancer morbidity and mortality, and the study of this spatial variation is one way of obtaining clues for identifying the relevant complex of environmental factors.

By the late 1960s, two major directions were appearing in the study of environmental carcinogenesis. The first was the potential impact of low level chemical

exposures on cancer patterns. The second was the recognition of the many factors comprising lifestyle, notably tobacco use and diet. There are two major categories of environmental exposure to carcinogens. "Lifestyle" includes factors that the individual himself can control such as smoking, alcohol, food, drugs, cosmetics, sunlight, and reproductive behavior. "General environmental factors" means exposure over which the individual has little or no control, such as occupational exposures, air and water pollution, food additives, biological contaminants, and radiation (Task Force on Environmental Health 1984).

The major conclusions of these studies are as follows. Herron (1969) showed that skin cancer patterns in Australia (like the U.S.) resembled the patterns of sunlight intensity and thus suggested a link between exposure to sunlight and skin cancer risk. Haber and Lipkovic (1970) studied the incidence of thyroid cancer in Hawaii in relation to environmental factors. The general conclusion was that the higher incidence of thyroid cancer was related to ionizing radiation. By studying the spatial variation of gastric cancer, Jakab et al. (1971) could link the causes of cancer incidence to the climate, crop, food, drinking water, and life-styles of a region. Hoover and Fraumeni (1975), investigating cancer mortality in the U.S. counties with chemical industries, found that geographic analysis of the U.S. cancer mortality (from 1950 to 1969) revealed excessive rates for bladder, lung, liver, and certain other cancers among males in 139 counties where the chemical industries are most highly concentrated. MacDonald (1976) studied the demographic variation in cancer in relation to industrial and environmental influence. He showed that the environmental factors of exposure over time to air and industrial pollutants

in Houston had a demonstrable effect in increasing regional mortality from cancer of the respiratory tract and heart disease.

A cartographic analysis of cancer mortality in the British Isles by Howe (1981) identified the urban affinity of lung cancer. Also, a study by Minowa et al. (1981) of the geographic distributions of lung cancer mortality and environmental factors in Japan revealed that standardized mortality ratios were higher in areas related to strong urbanization, coal and/or lignite mines, and steam power plants than in other areas. Greenberg (1983) suggested that the spatial convergence of cancer mortality in the U.S. and industrialized countries was caused by change in the geography of risk factors (air and water pollution, cigarette smoking, alcohol consumption, diet, occupation, and socioeconomic status, stress, and medical practices) associated with the diffusion of urban culture. Shannon and Pyle (1992) published a medical atlas for the 20th century to describe the geographic patterns of cancer mortality rates and environmental factors in the U.S. The atlas covers the changing patterns of disease (cancer) and medical care with reference to demographics, economics, geography, epidemics, and the supply and distribution of physicians and hospitals.

Geographical studies of disease are valuable for two major reasons. The first is that they suggest possible causal factors in pathogenesis. The second is that spatial patterns of disease may serve as useful indicators of how regions are structured, and of how individuals and groups exist in mutual interaction with the environment (Mayer 1983).

In summary, previous studies have suggested many factors which generate spatial variations of disease occurrence:

1. Physical environment - These elements may include factors of soil, air, water, climate, and geographic location; pollutants in the soil, air, and water.

2. Social, economic, and cultural factors - These are linked strongly to individual and collective behavior, such as smoking and diet.

3. Genetic factors - Many diseases are either genetically-based, or may be related to genetic predisposition.

The geography of disease may be viewed as resolution from numerous social, cultural, behavioral, environmental and biological factors operating either together or individually. Therefore, analysis of geographical patterns of disease and the social, cultural, economic, environmental and biological factors could be useful.

2.2 Spatial Analytical Methods for Disease Patterns

2.2.1 Methodological Progress for Analyzing Disease Patterns

One promising approach to studying the relationships of human, disease, and environment involves the use of spatial analysis in analyzing disease patterns. Spatial analytical techniques include disease-mapping and associative analysis.

The first disease maps appeared in the late 18th and early 19th century (Mead et al. 1988). These disease maps are a simple description of spatial arrangement of the data, with very little analysis provided. The most famous 19th century disease map was John Snow's dot map of cholera around the Broad Street water pump in London (1855). The map showed that cholera was a water-borne disease, with the pump the local source of infection. In the 1950s and 1960s, an increasing need emerged for statistical mapping that could rationalize the search for patterns and relationships.

Statistical mapping identifying a theoretical distribution of a disease makes use of a normal distribution. From this normal distribution, the probability of obtaining values larger than or equal to a certain value can be read, thus determining the areas that are least likely (significantly low) and most likely (significantly high) to have occurred by chance. Unlike normal probability distribution, Poisson probability distribution describes the likelihood of the occurrence of rare, random events, given a mean expectation and a variance. Choynowski (1959) used it to calculate the probability of any given incidence level in southern Poland. In comparing this probability map with an original map of mortality rates of brain cancer, he attempted to explain differences in terms of environmental factors. The technique of probability mapping of mortality has since been widely used, including applications of White (1972) and McGlashan (1972).

Standard death rates can be adjusted not only for differences in the proportion of the population at various ages but also for different proportions of sex, ethnicity, income, or other classifications. Howe (1960) first used the age-standardized mortality rates and mapped the geographical distribution of cancer mortality in Wales to identify possible environmental factors. Chiang (1961) later suggested in an U.S. Vital Statistics Report that standard error of mortality rates can be used to assess the significance of a particular rate. The standard error of a rate is a measure of the sampling variability of the rate. Armstrong (1969) proposed the use of standard deviation of the distribution of rates for the purpose of dividing the rates into mapping categories.

A cartogram is a map on which statistical information is presented in diagrammatic form. In a cartogram, some of the usual geographic qualities (such as size, shape, or contiguity) on the map are ignored, so that the areal units can be transformed to be

proportional to some other quality. Forster (1966) suggested the use of cartograms in which the size of the areal units on a map would be proportional to their at-risk population rather than to their land area.

Computer-assisted methods can promote the utility of mapping, data preparation, and analysis. McGlashan and Bone (1967) first used computer-aided methods in mapping mortality, and furthermore, Hopps (1969), a pathologist, used computer-aided contour maps in studying both mortality and environmental factors. Pyle's (1971) maps of mortality in the Chicago region used contour and trend-surface mapping. Among the more important examples of these types of maps were Learmonth's (1972) map of disease mortality in Australia and Armstrong's (1972) work dealing with mortality on the island of Hawaii. Mortality maps can suggest possible risk factors and identify areas in which epidemiologic and etiologic studies might be most rewarding. By viewing maps, information can be acquired about patterns and relations of spatial objects. One of the disadvantages of maps is that it is difficult to interpret them in an objective way.

A wide variety of techniques are available to determine the significance of disease distributions and methodological frameworks range from simple chi-square testing for significance to the use of complex multivariate explanations. It is a common practice in many health-related studies to use a chi-square structure in testing a hypothesis when the data distributions are normal (Griffiths 1971). Griffiths (1971) used chi-square testing and Kendall's rank in a nonparametric examination of different mortality levels in relation to social class. He discovered that in spite of much less manufacturing activity than many other British settlements, significant health status differences exist from one ward of

Exeter to another. Part of this difference was attributed to housing conditions as well as social class.

Given basic assumptions underlying certain principles in spatial analysis, there is little doubt that parametric testing procedures ranging from simple correlations through different kinds of factor analysis to multiple regression assist in explaining patterns. One method for testing distributions over time (how much variation over time exists) is simple linear regression. Multiple regression is used to examine the relationship between several independent variables and a single dependent variable (Blot et al. 1979). Cliff and Haggett (1988) used the ordinary least squares regression model and the residuals from this regression to demonstrate the interplay of two factors, defective drainage and source of water supply, in accounting for the geographical distribution of cholera deaths in the mid 19th century London. In particular, the mapping of the spatial distribution of the residuals can give an indication for associative relations.

As a next step, it is useful to examine various kinds of cancer along with other leading causes of death with multivariate methods. Factor analysis is a useful descriptive tool in such comparisons. Factor analyses are used to show disease clusters and can add some explanation to the differences in levels of significance when comparing cancer distributions. The first factor analytic approach in understanding disease distributions within Chicago was accomplished in 1968 (Pyle). The results indicated that population density has a stronger relationship with some contagious diseases than the poverty-health syndrome indicators. There have been factor analysis studies by Inaba et al. (1981), Dent et al. (1982), Lam (1986), and Babin (1986).

Canonical correlation analysis measures the relationships between a large number of independent variables and a large number of dependent variables. The method was used in the Akron, Ohio area by Pyle and Lauer (1975) and in Houston, Texas area by Briggs and Leonard (1977). In these analyses, they found that sets of disease groups have strong associations with indicators of a poverty syndrome in health status. More recently, Babin (1986) used Z scores, factor, and canonical correlation for cancer mortality distribution in the U.S. Male stomach cancer and female lung cancer developed the most distinctive spatial patterns and lung cancer was found to be associated with the distributional patterns of some occupational environmental variables for both male groups.

As different approaches, spatial autocorrelation and trend surface analysis were tried by Glick (1977, 1979b, 1982). Spatial autocorrelation measures the association between a disease's values taken from two adjacent places. Trend surface analysis, a special case of multiple regression, can be used to examine the geographic distribution of diseases by finding the best-fit surface. Furthermore, the examination of residuals can suggest variables to influence the data distribution. A result of these analyses was that the spatial distribution of cancer mortality among the counties in Pennsylvania is significantly clustered for lung, stomach, and bladder cancer, while it is statistically random for breast and cervical cancer and leukemia. Kennedy (1988), using the same method, concluded that male lung cancer exhibits spatial autocorrelation while female lung cancer does not, and that the female data exhibit a spatial trend while the male data do not.

Discriminant analysis is useful to analyze the differences between the groups and gives a means to classify any case into the group which it most closely resembles. A stepwise discriminant analysis was used to identify cancer types that were critical to the

formation of regions (Babin 1979). Babin indicated that cancer mortality from 1950 to 1969 in the U.S., the South, the Northeast, and the West emerged several times as areas of distinctive patterns (liver, prostate, non-melanoma skin, rectal, and esophageal cancer).

A disease cluster occurs when the number of disease cases within a geographic area, a particular group of people, or a certain period of time is greater than expected. Clustering analysis is a categorization technique that is based on the "taxonomic distance" between data items. McGlashan (1983) demonstrated how this technique can be used to group both cancer mortality (R-mode) and areas where the disease occurred (Q-mode) among goldminers in several African countries. Disease junctions in the cluster dendograms produced suggested similar causation, while areal junctions suggested common environmental influences such as climate.

Spatial (temporal, space-time) scan statistic has been recently developed for detecting clustering (Kulldorff and Nagarwalla 1995; Kulldorff 1997). It searches for clusters of cases without specifying their size or location ahead of time and tests for their statistical significance while adjusting for the multiple testing inherent in such a procedure. Hjalmars et al. (1996) studied the childhood leukemia in Sweden, using GIS and a spatial scan statistics for cluster detection. They found no significant clusters in leukemia in Sweden, unlike earlier studies (Openshaw et al. 1988; Knox and Gilman 1992). Kulldorff et al. (1997) employed a scan statistic for breast cancer clusters in the northeastern U.S. The study showed that there is a statistically significant and geographically broad cluster of breast cancer deaths in the New York City-Philadelphia, Pennsylvania, metropolitan area. In addition to this, there are many different statistical tests for disease clustering such as descriptive cluster detection (Openshaw et al. 1987),

cluster detection (Turnbull et al. 1990), focused (Bithell 1995), global clustering (Besag and Newell 1991), and space-time interaction (Kulldorff and Hjalmars 1999) method.

A fractal is a spatial set that manifests a regular scaling relationship between the number of its constituent elements and their measure (size, density, and so forth) (Lam and Cola 1993). In a fractal analysis of disease distribution, Lam et al. (1993) demonstrated that distinct self-similarity ranges exist in the three leading cancer mortality patterns (stomach, esophagus, and liver) in China.

Artificial neural networks have been developed as generalizations of mathematical models of human cognition or brain biology (Fausett 1994). Recently, the integration of artificial neural networks with GIS has been applied to the spatial distribution of epidemic disease (Feng 1998). Feng explored the use of artificial neural network for the projection of the AIDS incidence rates in the U.S. He showed that both projected and back projected AIDS cumulative incidence rates are close to their reported values, suggesting that the neural networks are successful for estimating the AIDS epidemic in the country for the period 1982-1993.

2.2.2 Geographic Information System (GIS) and Medical Geography

The importance of GIS within medical geography has risen sharply over the last few years. There is much potential for the use of GIS in the management and analysis of geographic pattern of disease and environmental factors. Geographic information system (GIS) can be defined as a computer-assisted information management system of geo-referenced data. The first GIS was developed in North America after R.F. Tomlinson proposed the basic concept of a GIS in the 1960s. This system integrates the acquisition, storage, analysis, and display of geographic data. The application fields and objectives of

a GIS can be varied and they concern a great number of questions linking social and physical problems such as transport and agricultural planning, environment and natural resources management, location/allocation decisions, facilities and service planning (education, police, water, and sanitation), and marketing.

Medical geography, which is the geographical study of human health, including the provision of health care, has been studied since the 18th century. It has developed rapidly as a subfield in geography since J.M. May's study in 1950. It covers the description of the spatial distribution of disease and health care and considers possible causative relationships among sickness, disease, and death and local variations in environmental conditions. A more recent interest is on the spatial aspects of organization of health services, especially the identification of optimum locations for health care facilities (Goodall 1987). Furthermore, medical geography may play an important role in the improvement of health, healthy environment, and medical planning.

GIS can assist in health research, in health education, and in the planning, monitoring, and evaluation of health programs. GIS has been rapidly developed since the 1980s and its potential in medical geography has been recognized by many scholars (Lam 1986; Openshaw et al. 1987, 1988; Twigg 1990; Scholten and de Lepper 1991; Verhasselt 1993; Loslier 1994; Douven and Scholten 1995). They supported the idea of using a GIS approach because of its capability and potential value to 1) overlay and integrate spatial information, 2) substantiate quantitative analyses in disease ecology and health care delivery through its capability to handle large amounts of data (Verhasselt 1993), and 3) improve public health and understand environmental risk (Scholten and de Lepper 1991).

Although GIS is important in linking data to geographic locations, displaying spatial information, conducting spatial analyses, and producing detailed maps about health care and illness, the roles of GIS in medical geography include alternative modeling of health care resources and making widely integrated studies possible. In other words, GIS is capable of integrating data from various sources to provide information for effective and correct decision-making for medical planning. For example, GIS can integrate and measure various layers of environmental data (such as climate, terrain, altitude, soil, vegetation, land use, and hydrology) and real time data. In investigating spatial distribution of health care resources, GIS can store not only the length and width of roads and buildings but also the non-spatial data. It can study temporal and spatial change in these data. Buffer zones can be generated to investigate illness at or near pollution source and other hazardous sites.

Recent developments in GIS have made it possible to analyze complicated spatial relationships effectively. Spatial autocorrelation analysis that requires GIS for distance and centroid calculations has been used (Kehris 1990; Fan et al. 1993; Chou 1993). The statistical algorithms of SaTScan (*http://dcp.nci.nih.gov/BB/SaTScan.html)* and SpaceSat (*http://www.spacesat.com*) have been integrated into GIS software. Existing GIS's capabilities for more sophisticated forms of spatial analysis and decision making, however, are rather limited. This lack of integration of spatial statistical procedures and spatial models is perceived as a major shortcoming (Openshaw 1990; Goodchild 1992; Anselin and Getis 1992; Fischer and Nijkamp 1992). Moreover, there are some problems concerning data collection and verification in GIS health care research, because data sets

tend to have been compiled for large administrative units (Twigg 1990) and have come from different organizations in a variety of incompatible spatial units (Heywood 1990).

Most research on the potential of GIS in medical geography has been accomplished in spatial distribution based on geocoding systems and computerized mapping systems. In cancer-related research, Openshaw et al. (1987, 1988) developed a geographical analysis machine (GAM) and used this facility to investigate the association between proximity to nuclear power stations and childhood leukemia cancer clusters in the northern region of England. Within the GIS component of GAM, a grid of points was superimposed over the study area. Each point of the grid was used as a center. The age-sex adjusted incidence rates of childhood leukemia were calculated for each circle and tested for significance based on Poisson probabilities. Fitspatrick-Lins et al. (1990), using exploratory data analysis (EDA) and GIS, found that radon potential was high for the Piedmont Upland of Fairfax County, Virginia. Moore (1991) used PC ArcInfo for air toxics risk assessment for emissions (carcinogenic and noncarcinogenic) from facilities in California. In one part of the research, isopleths of cancer risk were drawn around a facility. Kulldorff et al. (1997) employed a recently developed spatial (temporal, space-time) scan statistic to analyze breast cancer clusters in the Northeast U.S. This study used case/mortality and population counts for a set of census areas, the geographical coordinates (latitude and longitude) for each of those areas, and covariates (such as age, sex, race, urbanicity, and/or parity). In addition to these studies, other examples include: using GIS to evaluate radon potential and housing issues in Florida by Fandrich and Zwick (1993), developing maps of radon potential for the state of Florida by Zwick and Latiner (1993), evaluating clusters of adverse health outcomes (cancer) by Aldrich et al. (1994),

GIS tracks of Long Island breast cancer by a New York State Health Department (1994), and modelling the possible association between larynx cancer and waste incineration by Gatrell et al. (1995). Also, a few health services are in the process of conducting cancer research on the potential of GIS. After the collection and integration of data, mapping and map overlay, and statistical analysis, it is possible for GIS to monitor the phenomena as well as provide analytical capability to model or forecast for medical planning.

Based on studies such as the above mentioned, it is expected that the application of GIS in medical geography will accelerate the search for the relationship between cancers and environments, the causes of cancer, and the development of medical geography.

2.3 Geographical Distribution of Cancer and Environment

2.3.1 Cancer Mortality Trends from International and National Perspectives

Cancer occurs throughout the world. Cancer has been defined as uncontrolled new growth which invades and destroys living tissue (Shannon and Pyle 1992). The cancer mortality rate is the number of deaths with cancer given as the underlying cause of death occurring in a specified population during a certain time period. It does not necessarily coincide with the incidence of new cases of cancer in that year.

International cancer mortality trends

World patterns of cancer mortality and incidence reflect regional variations in exposure to not only hazards of the environment but also racial and cultural characteristics of the people living in different parts of the world. Dunham and Bailar (1968) produced 16 world maps showing the distribution of various types of cancers. International patterns of cancer have been carefully analyzed with reference to cancer

etiology (Muir and Nectoux 1982; Waterhouse et al. 1982; Page et al. 1985; Howe 1986b; Muir et al. 1987; Davis and Hoel 1990; Higginson et al. 1992).

In general, cancers of the lung, trachea, and bronchus are the common cancers in most Western countries. Not only are these types common, but their incidence continues to increase. In particular, lung cancer death rates are highest in Scotland and the Netherlands and are relatively low in Chile, Ecuador (and in other Latin American countries), Thailand, and most African countries (Page and Asire 1985; American Cancer Society 1993). Lung cancer is not only a major cause of death in the U.S. but is also one of the three most common cancers (Lung, Prostate, and Stomach) in men throughout the world. Since 1987, lung cancer in the U.S. has been overtaking breast cancer as the most common cause of cancer death among women.

Cancers of the colon and rectum are often considered together as colorectal cancers. Colorectal cancers are considered diseases of economically developed countries. They are common in Europe, North America, and New Zealand but rare in Asia and uncommon in tropical Africa and most Latin American countries (Waterhouse et al. 1982). In the U.S., deaths from colorectal cancer rank second only to those from lung cancer in the total number of deaths, but these death rates vary widely by geographic area in this country.

Prostate cancer death rates are high among northern Europeans such as Swedes, Norwegians, and Swiss but rare in East Asia. In the U.S., cancer of the prostate is the second most common cause of cancer death among men. Female breast cancer is rare in much of Asia, but it has been one of the most common causes of cancer death in North America and most of Europe (Page and Asire 1985).

Stomach cancer develops frequently in Japan, Korea, China, and the Soviet Union but not in Thailand. Also the death rates are relatively high for both men and women in Chile, Hungary, Austria, Germany, and Poland. They are significantly lower, for both sexes, in Israel, Canada, and Australia. Americans have stomach cancer death rates among the lowest in the world.

In addition to these cancer sites, cervical cancer is a common form of cancer among women in much of Latin America and the Caribbean. Esophagus cancer is characterized by wider variations in incidence and death rates than any other cancers. Esophagus cancer death rates are very high in Hong Kong, Chile, and parts of the former Soviet Union but low in Israel. Leukemia cancer death rates are very high in Denmark, Italy, and the U.S., whereas they are generally low in Southeast Asian nations such as Thailand, Hong Kong, and the Republic of Korea. Liver cancer death rates are high among the indigenous inhabitants of tropical Africa, in China and other parts of the Far East, in the Western Pacific, and in South or West Asia, but they are rare in countries such as Canada or the United Kingdom.

In developed countries, cancers are mostly associated with old age and strike mainly the lung, female breast, prostate, colorectum, and cervix, whereas in developing countries the most common cancer sites are the cervix, mouth-pharynx, liver, esophagus, and stomach (except for Japan) afflicting people in their thirties and forties (World health Organization 1987-1990).

National cancer mortality rates and trends

Many studies have given a detailed picture of cancer mortality and cancer incidence in the U.S. (Burbank 1971; Lilienfeld et al. 1972; Devesa and Silverman

1978; Pollack and Horm 1980; Greenberg 1983; Shannon and Pyle 1992). The National Cancer Institute has published atlases of cancer mortality by county (Mason et al. 1975, 1976; Riggan et al. 1987; Pickle et al. 1987, 1990) and cancer statistics (Mason et al. 1974; Riggan et al. 1983; National Cancer Institute 1988; Miller et al. 1993; Howe and Lehnherr 1996). The American Cancer Society has published "Cancer Factors and Figures" for every year since 1985 (American Cancer Society 1985-1999).

There has been a steady rise in the cancer mortality rate in the U.S. in the last half-century. The age-adjusted rate in 1930 was 143 per 100,000 population. It rose to 152 by 1940, to 157 in 1950, to 163 in 1970, and was 174 in 1990 (American Cancer Society 1993, 1994, and 1995). The major cause of this increase has been lung cancer. If lung cancer deaths were excluded, cancer mortality would have declined by over 14 percent between 1950 and 1990 instead of increasing by 10 percent (Miller et al. 1993). Mortality rates for all cancers and all cancers excluding lung and bronchus were highest in black men and lowest in white women.

The dominant cancer sites of mortality in the U.S. are lung, colon and rectum, male prostate, female breast, and pancreas (Figures 2.1 and 2.2). Lung cancer death rates have been climbing sharply since the 1930s for men and since the early 1960s for women. Breast cancer had been the major cause of cancer deaths among all the U.S. women from 1945 to 1988. Death rates from colorectal cancer and prostate cancer both peaked in the late 1940s, then dropped and leveled off. The nation had high stomach death rates in the past, from 1930 to 1940, but stomach death rates have been declining markedly since 1940. Death rates from cancer of pancreas rose slightly from the 1930s into the 1960s and then leveled off. Until about 1945, uterus cancer was the major

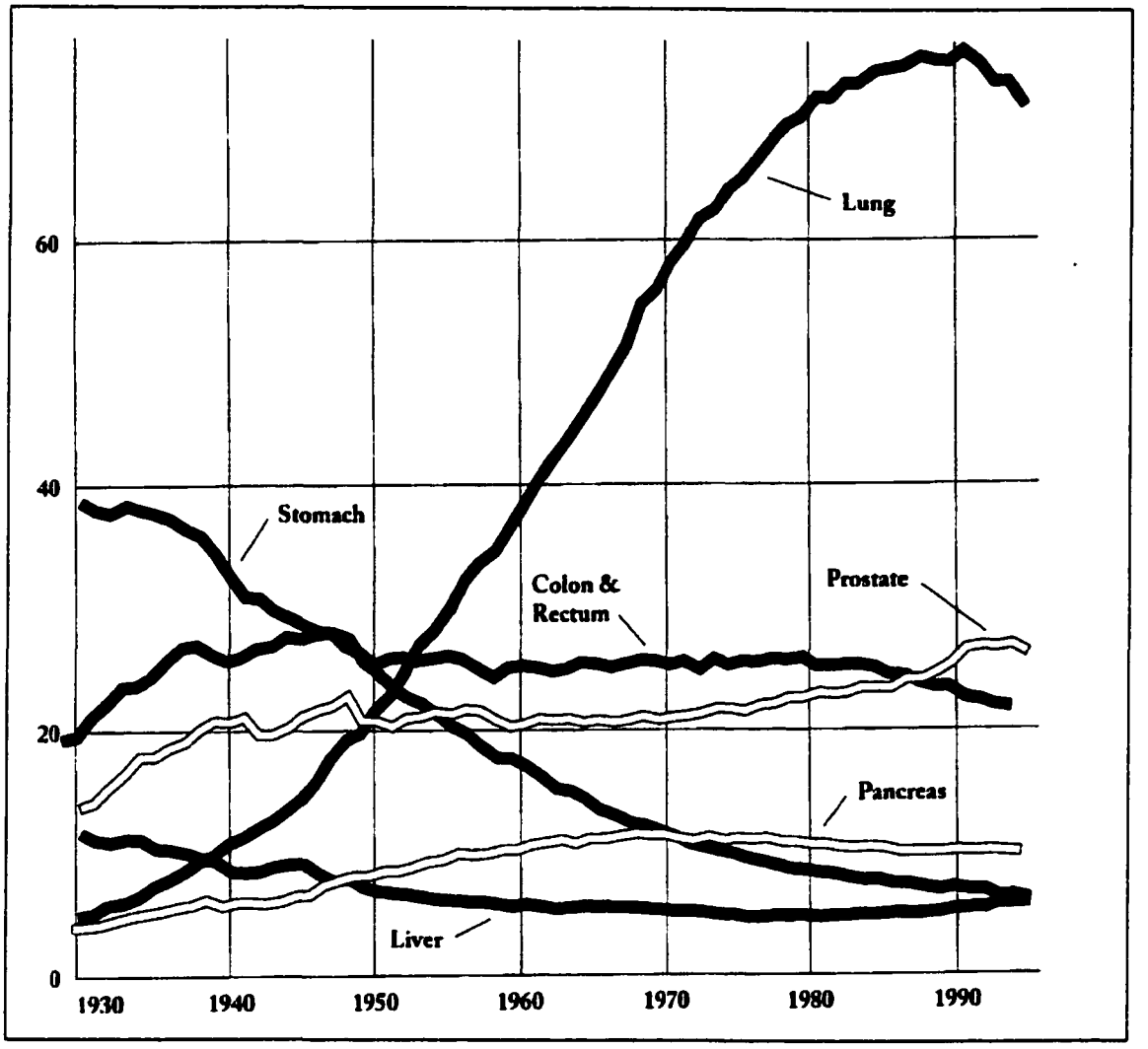

Figure 2.1 Age-adjusted Cancer Death Rates, Males by Site, U.S., 1930-1994
Rates are per 100,000 and are age-adjusted to the 1970 U.S. standard population.
(Source: Vital Statistics of the U.S., 1997)

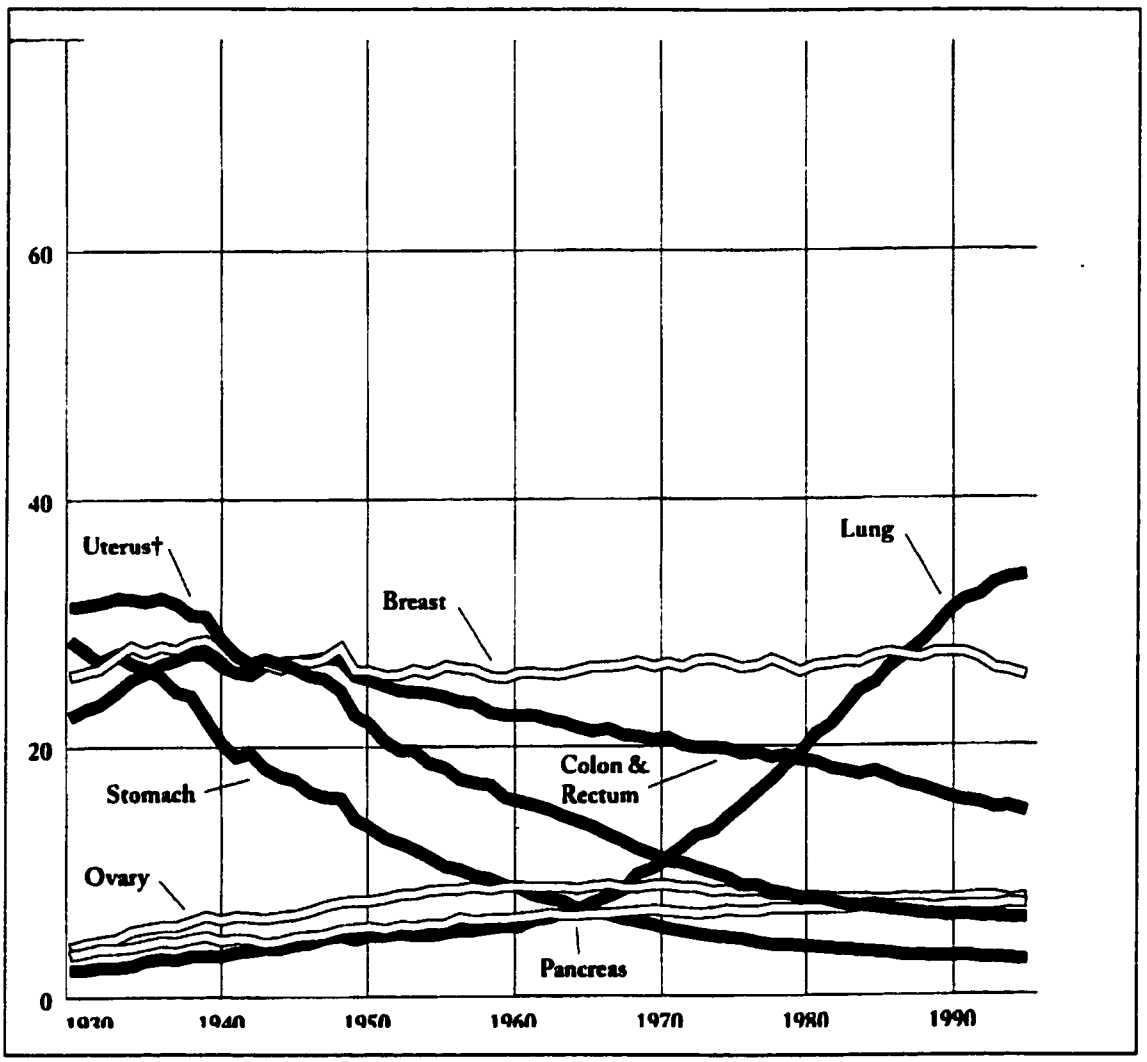

Figure 2.2 Age-adjusted Cancer Death Rates, Females by Site, U.S., 1930-1994
Rates are per 100,000 and are age-adjusted to the 1970 U.S. standard population.
(Source: Vital Statistics of the U.S., 1997)

cause of cancer death among all American women, but deaths from this cancer have been declining steadily since 1935. Esophagus and bladder cancer death rates for men have remained almost constant since 1930. Death rates from cancer of leukemia increased slightly from the 1930s into the 1960s and then decreased (American Cancer Society 1994).

The age-adjusted incidence rates for all cancers and both sexes combined increased in both blacks and whites by about 19 percent between 1973 and 1990 (Miller et al. 1993). However, the magnitude of the overall cancer incidence rate among blacks was nearly 8 percent higher than whites in 1990. Incidence rates for all cancers and all cancers excluding lung and bronchus were highest in black men and lowest in black women among all race and sex groups.

According to studies by Greenberg (1983), Pickle et al. (1987 and 1990), and Shannon and Pyle (1992), mortality rates for many cancers were higher in the Northeast, in particular, urban areas.

Cancer mortality rates of trachea, bronchus, lung, and pleura were high in urban areas of the North and were clustered especially along the Southeast Atlantic and Gulf coasts (in particular, Louisiana).

Cancer mortality rates from the colon and rectum usually varied widely by geographic area. The death rates were highest, for example, in Northeast urban areas and lowest in the South and Southwest.

Most states in the urban Northeast and in several metropolitan areas had rates of breast cancer mortality that were near the average or somewhat above. The parts of the country with the lowest mortality of breast cancer included large parts of the Hispanic

Southwest (excluding southern California). The breast cancer mortality rates for Chinese and Japanese women in the U.S. were three times higher than in Singapore or Japan, but not so high as the death rates among whites in the U.S. The breast cancer death rates among American black women were lower than among white women, but are increasing.

Prostate cancer, the second most common type among men, showed the highest death rates in Montana, Dakota, Mississippi, South Carolina, Wyoming, while some of the lowest rates showed up in Rhode Island, New York, and Massachusetts. For pancreas cancer mortality rates among men, a clustering of high rate areas could be detected in the urban Northeast and the South Louisiana. Cancer mortality rates of the oral cavity (for male) and esophagus were high in the urban Northeast and in several metropolitan areas.

In particular, death rates from stomach cancer were high until the 1930s but they have since been declining steadily. Stomach cancer mortality rates of blacks have been consistently high in New York City and southern Louisiana, whereas those of whites have featured an excessive cluster in rural counties in the North Central region and northern urban areas (Pickle et al. 1987 and 1990).

2.3.2 Cancer and Environment in Louisiana

Excess of cancer morality rates for the oral cavity and the respiratory system has been observed in Louisiana since the early 1930s (Gover 1940). From the 1950s to the 1970s, significantly higher mortality rates were observed for all cancers combined among men in South Louisiana (Riggan et al. 1983). All four sex-race groups (males and females for whites and nonwhites) in Louisiana have had higher mortality rates for

all cancers combined when compared to the national averages since the 1980s (Miller et al. 1993). Cancers that are excessive in Louisiana compared with other states are cancers of the lung, prostate, larynx, mouth, and esophagus. The increase in all cancer sites combined is primarily due to an excess in lung cancer (Riggan et al. 1983; Miller et al. 1993; Groves et al. 1996). It accounts for 85% of the excess of cancer over the national average in Louisiana (American Cancer Society 1990a).

In general, cancer mortality rates in Louisiana are high in New Orleans, Iberia, Plaquemines, St. Martin, Acadia, St. Bernard, and St. Mary parishes and relatively low in the parishes of East Feliciana, St. Helena, Sabine, Caldwell, and Red River (Figure 2.3). The higher rates of cancer mortality for lung and other sites are more prominent in the southern part of the state, but rates have also recently increased significantly in the northern part of the state and on the west bank of the Mississippi River (Figure 2.4). For the same cancer, females always have lower mortality than males, and the spatial patterns between males and females for the same site generally exhibit the same trend.

A majority of human cancer may be caused by chemical carcinogens in the environment (Higginson 1993). Many substances are classed either as human carcinogens (Table 2.1), which show sufficient evidence for carcinogenicity in humans, or as possible human carcinogens (Table 2.2). Table 2.3 includes those substances which were released in Louisiana, as reported by the Environmental Protection Agency (EPA). Even though Louisiana toxic chemical releases and transfers from 1987 to 1992 have decreased, Louisiana has long been one of the leading states in the amount of carcinogenic and non-carcinogenic Toxics Release Inventory (TRI) releases and transfers (Figures 2.5 and 2.6). In particular, South Louisiana has shown the highest

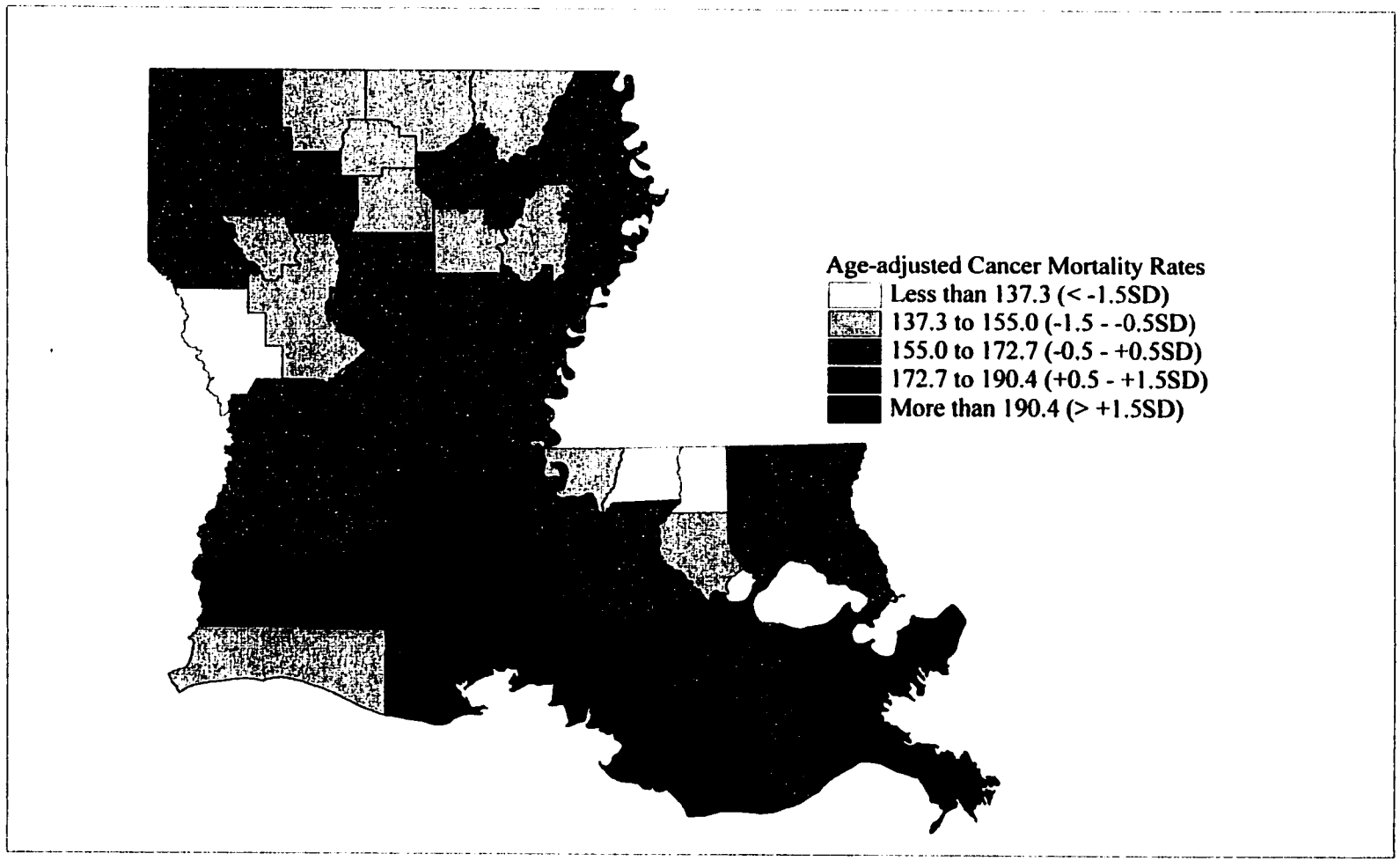

Figure 2.3 Cancer Mortality Rates for All Sites, Louisiana: 1953-1987
(Source: Calculated by Author from National Technical Information Service, 1992)

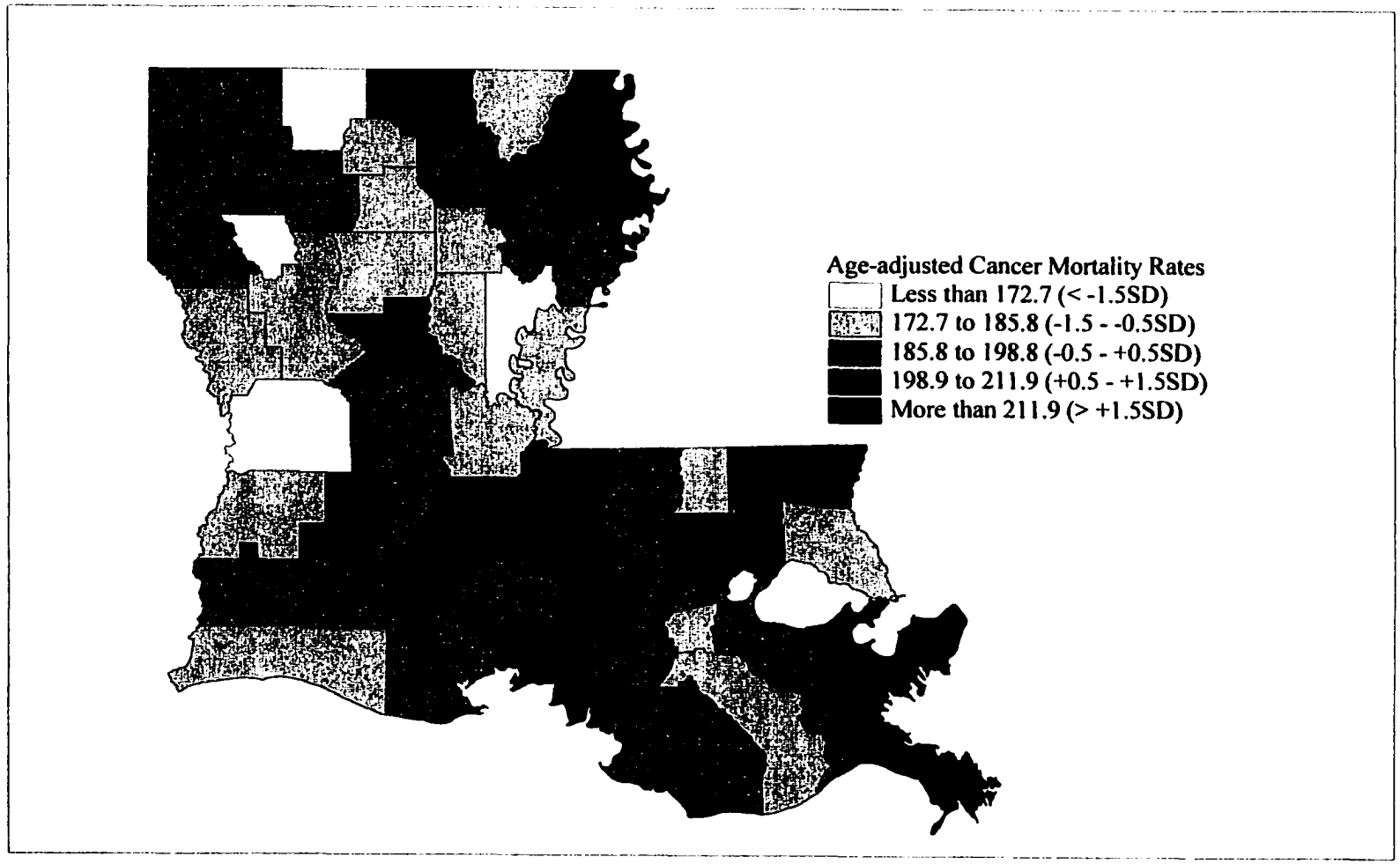

Figure 2.4 Cancer Mortality Rates for All Sites, Louisiana: 1988-1996
(Source: Centers for Disease Control and Prevention (CDC) Wonder, 1999)

Table 2.1 Substances Classed as Human Carcinogens

Carcinogens	Main Potential Source of Exposure	Site of Cancer
• **4-Aminobiphenyl**	Dye manufacture	***Bladder***
• **Arsenic and arsenic compounds**	Production and use of arsenical insecticides, mining, copper smelting	***Skin, lung, liver***
• **Asbestos**	Mining, insulation material production and use, textile shipbuilding, brake lining(repairing)	***Lung, pleural & peritoneal mesothelioma***
*Indoor pollution: roofing and flooring materials, textiles, papers, filters and gaskets, cement pipes, coating materials, thermal/acoustic insulation		
• **Auramine manufacture**		***Bladder***
• **Benzene**	Rubber in industry, glues industry, shoe, dye manufacture, petroleum refining	***Leukemia, bone marrow***
• **Benzidine**	Dye manufacture	***Bladder***
• **Bis(chloromethyl)ether and chloromethyl methyl ether(technical grade)**	Ion-exchange resins, chemical industry	***Lung***
• **Chromium compounds, hexavalent**	Chromate pigments production and use chromium plating and alloy production, stainless steel welding	***Lung***
• **Coal-tar pitches, coal-tars, PAHs, bitumens**	Coal gasification, coal distillation, coke production	***Lung, oral cavity skin, bladder***
• **Ethanol**	Alcohol	***Oral cavity***
• **Formaldehyde**	Urea for maldehyde insulation, plywood, particle board, floor coverings, cosmetics	***Nose, nasopharynx***
• **Wood dust**	Furniture and cabinet-making	***Nose***
• **Hematite mining, underground with exposure to radon**	Mining	***Lung***
• **Isopropyl alcohol manufacture(strong acid process)**	Isopropylene manufacturing	***Nose***
• **Mustard gas (sulphur mustard)**	Poison gas production	***Lung, trachea, bronchi, larynx***
• **2-Naphthylamine**	Dye manufacture	***Bladder***
• **Nickel (compounds)**	Nickel refining	***Nose, lung***
• **Radon**	Concrete and brick building materials; natural gas used in home	***Lung***
• **Shale-oils**	Shale oil industry	***Skin***
• **Soots**	Chimney sweep	***Skin***
• **Tobacco smoke (polycyclic aromatic hydrocarbons)**	Personal and passive smoking	***Lung***
• **Vinyl chloride**	PVC & plastics manufacturing, polymerization	***Liver angiosarcoma***

(Source: Modified from Higginson et al., 1992 (pp.98-99 and p.78) and Task Force on Environment Health, 1984 (pp.12-13, pp.93-94, and pp.124-125))

Table 2.2 Substances Classed as Possible Human Carcinogens

Carcinogens	Main Potential Source of Exposure	Site of Cancer
• **Acrylonitrile**	Synthetic, fibers, resins, plastic manufacturing	***Lung***
• **Acetaldehyde**	Chemical intermediate, food additive, fruit and fish preservative	***Experimental evidence***
• **Aflatoxin**	Food additive(peanut derivatives)	***Liver***
• **Amitrole**	Herbicides	***Experimental evidence***
• **Beryllium**	Refining	***Lung, experimental evidence only***
• **Cadmium & cadmium compounds,**	Cadmium smelting, batteries production electroplating, cadmium alloy production	***Lung, prostate, kidney***
• **Carbon tetrachloride**	Production of fluoro-carbons, solvents, fumigants, pesticides	***Liver***
• **Creosotes**	Application as wood preservative	***Skin***
• **2,4-Diaminotoluene**	Manufacture of toluene-cyanates, dyes, hair dye formulation	***Experimental evidence***
• **1,2-Dichloroethane**	Intermediate in vinyl chloride production use as soil fumigant and solvent	***Experimental evidence***
• **1,1-Dimethylhydrazine**	Production	***Experimental evidence***
• **1,4-Dioxane**	Solvent, stabilizer in chlorinated solvents	***Experimental evidence***
• **Epichlorohydrin**	Production	***Lung***
• **Ethylene oxide**	Production, use as fumigant and sterilant	***Lymphatic & hematopoietic system***
• **Formaldehyde**	Production, manufacture of resins and, plastics use as disinfectant, fumigant and preservative	***Nose, nasopharynx***
• **Hexachlorobenzene**	Use as fungicide	***Experimental evidence***
• **Hydrazine**	Rocket fuels, herbicides, medicinals	***Experimental evidence***
• **o-Toluidine**	Intermediate in dye production	***Bladder***
• **Polychlorinated biphenyls**	Production, use in flame retardants, plasticizers, pesticides extenders	***Skin(melanoma), liver***
• **Tetrachloroethylene**	Dry cleaning, metal degreasing intermediate in fluorocarbon production	***Lymphatic & hematopoietic, urogenital system***

(Source: Modified from Higginson et al., 1992 (pp.100-106 and p.78) and Task Force on Environment Health, 1984 (pp.14-15, pp.93-94, and pp.124-125))

Table 2.3 Possible Human Carcinogens in Louisiana

Human Carcinogen	Source of Exposure
• Arsenic compounds (Arsenic trioxide is the only compound classified as human carcinogens)	Agricultural pesticide
• Asbestos (friable)	Chemical industry, cement manufacturing
• Benzene	Petroleum, petrochemical refinery, chemical, shoe, steel industry
• Chromium compounds	Petroleum refinery
• Vinyl chloride	Chemical industry

Probable Human Carcinogen	Source of Exposure
• 1,1-Dimethyl hydrazine	Chemical industry
• 1,2-Dichloroethane	Oil refinery, chemical industry,
• 1,4-Dioxane	Chemical industry
• 02-Nitropropane	----
• 2,4-Dinitrotoluene	Chemical industry
• Acetaldehyde	Chemical, petrochemical Industry,
• Acrylonitrile	Chemical industry
• Bis(2-chloroethyl) ether	Chemical industry
• Carbon tetrachloride	Chemical, petrochemical refinery agricultural chemicals
• Chloroform	Chemical, paper mill industry
• Creosote	Petrochemical industry
• Diaminotoluene (mixed isomers)	Chemical industry
• Epichlorohydrin	Chemical industry, oil refinery,
• Ethylene oxide	Chemical, petrochemical industry,
• Formaldehyde	Chemical, petrochemical, wood and forest products industry
• Hexachlorobenzene	Chemical industry
• Hydrazine	Oil refinery, chemical industry
• O-Toluidine	Chemicals industry
• Polychlorinated biphenyls(PCB's)	Agricultural chemicals (pesticide)
• Tetrachloroethylene	Chemical industry
• Trichloroethylene	Chemical industry, agricultural herbicide

(Source: Modified from Louisiana Toxics Release Inventory (1987-1992) and Corporate Release Challenge '92 (1993))

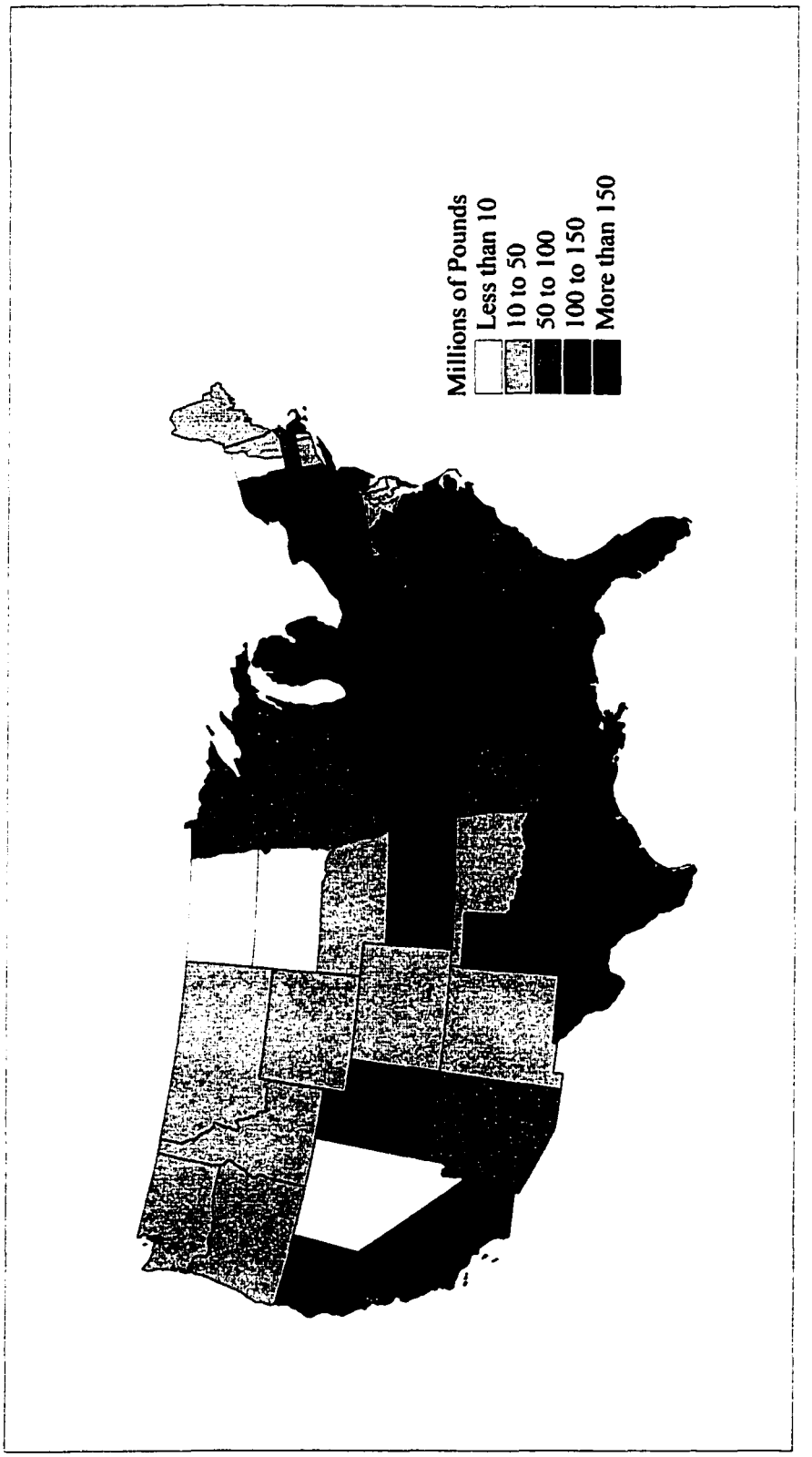

Figure 2.5 Toxic TRI Releases and Transfers, U.S.: 1990
(Source: U.S. EPA. Toxic in Community, 1992)

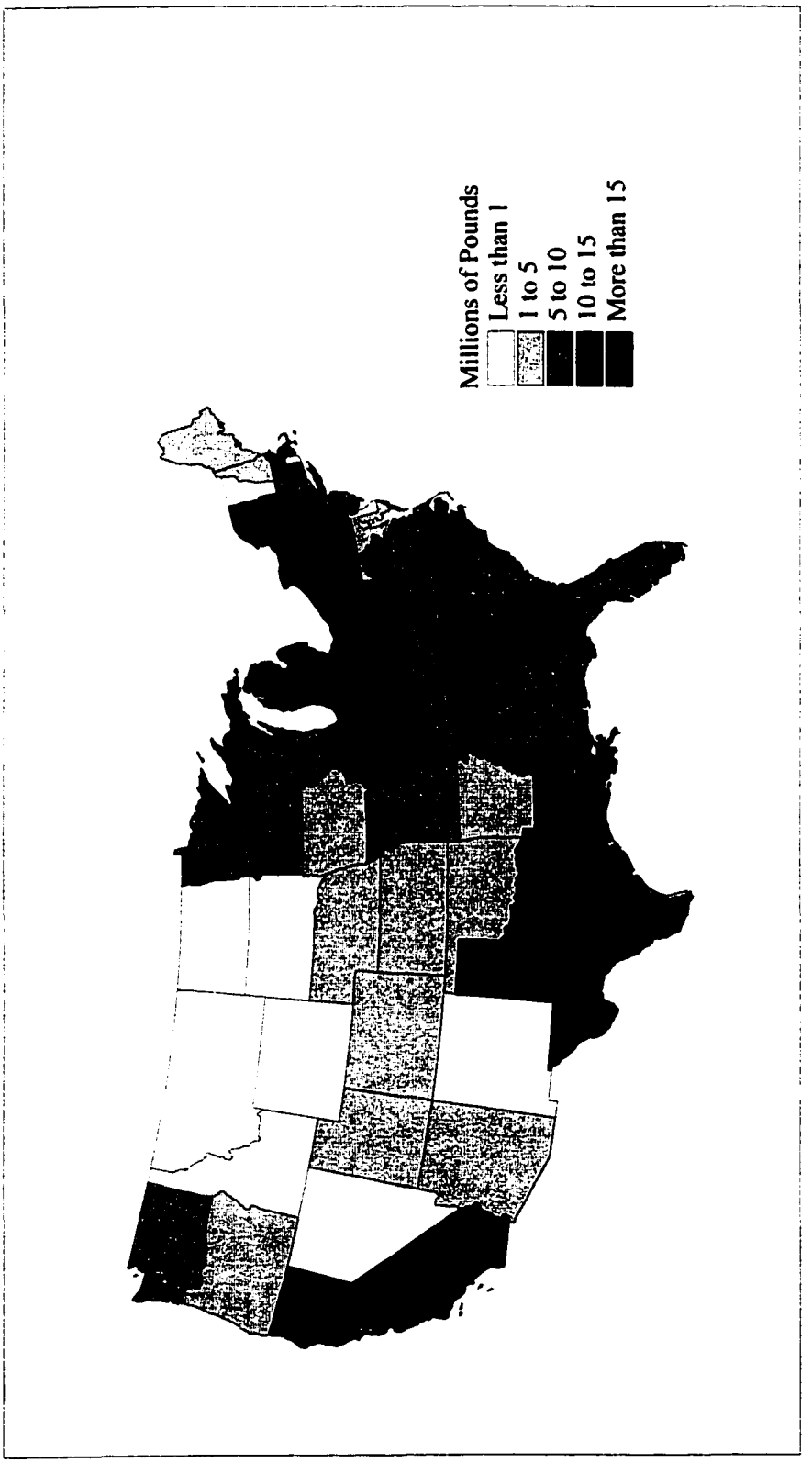

Figure 2.6 Carcinogenic TRI Releases and Transfers, U.S.: 1990
(Source: U.S. EPA. Toxic in Community, 1992)

amounts of total TRI releases (Figure 2.7) as well as carcinogenic TRI releases. In the 1990s, St. Charles, Ascension, Iberville, Ouachita, Calcasieu, and East Baton Rouge are among the highest in TRI releases of carcinogens, amounting to 90% of the state's water carcinogen TRI releases. A significant source of these pollutants has been attributed to industrial growth along the Mississippi River, discharges from river traffic, and northern point sources (Louisiana Department of Environmental Quality 1992).

Approximately 24% of Louisiana residents are classified as living below the poverty level, compared with 13% nationwide (Office of Planning and Budget 1994). The highest proportions of black population and of families below poverty level are concentrated in North Louisiana and especially along the west bank of the Mississippi River. Louisiana has a large proportion of indigent population, ranking second in poverty in the U.S. While basic medical care is available to indigents through the Charity Hospital system, cancer preventive practices are very limited (Chen et al. 1994).

The state has a culturally diverse population including French-Acadians, Anglo-Saxons, Scotch-Irish, and other non-French Europeans. South Louisiana is usually dominated by the "Cajun" culture and Roman Catholic (from French-Acadians), urban residents, and workers engaged in mineral-production-related jobs, which include the vast petroleum industry. But North Louisiana is dominated by Protestant culture, agriculture, and workers engaged in agriculture-related jobs. Geographically, South Louisiana generally consists of flat, low-lying areas, with many marshes and waterways while North Louisiana consists of hills, with few waterways.

Excess cancer mortality rates in South Louisiana, compared with the national averages, have been noted for a half-century. The Mary Bird Perkins Cancer Center

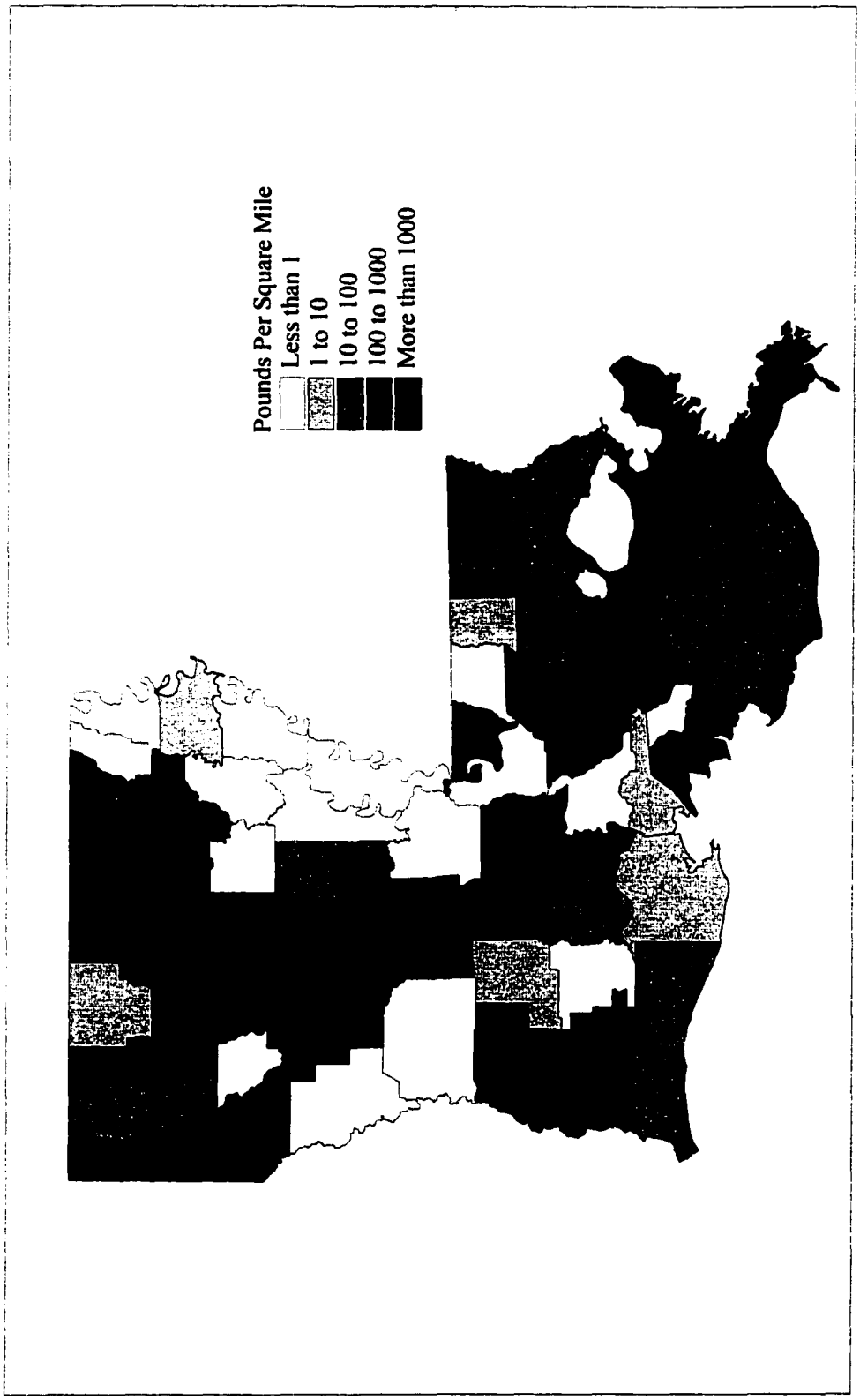

Figure 2.7 Total TRI Releases, Louisiana: 1990
(Source: Louisiana Department of Environmental Quality, 1993)

(1993) in Baton Rouge, LA reported that in South Louisiana 70-80% of cancer deaths are linked to lifestyle, environmental pollution accounts for less than 2%, and occupation accounts for only 4-6% of cancer deaths. Because of public concern of the high cancer mortality rates, the region's large petroleum and chemical industries, and the use of chlorinated water from the Mississippi River, cancer incidence rates in South Louisiana or the Mississippi River parishes have been studied (Chen et al. 1990, 1991, 1996, and 1998; Groves et al. 1996). However, cancer incidence in South Louisiana from 1983-1987 and Louisiana from1988-1992[2.1] is generally similar to or lower than SEER rates.

2.3.3 Studies on Cancer and Environment in Louisiana

A number of studies on the Louisiana environment-cancer question have already been carried out (Table 2.4). According to the report of ENSR Health Science, four types of cancer studies were identified: ecologic studies, case-control studies based on information obtained from death certificates, case control studies using information gathered from direct interviews, and occupational cohort studies (Wong and Foliart 1989, 1993). Ecologic study is a type of descriptive study based on grouped data in which the unit of analysis is populations and not individuals. Correlations based on aggregate data may mask a true association at the individual level and so such studies cannot assess a causal relationship. Case-control study is a type of analytical study which compares data on individuals with an illness (cases) with similar healthy individuals (controls). Cohort study is a type of analytical study in which the

[2.1] In 1990, cancer incidence rates (1983-1987) for South Louisiana excluded the New Orleans area and in 1996, the statewide cancer incidence rates (1988-1992) became available for the first time.

Table 2.4 Studies on Cancer and Environment in Louisiana

Author/Year	Content
Correa & Johnson (1983b)	Cancer and lifestyle
Correa et al.(1988)	Diet, nutrition, & cancer
Chen et al. (1988)	Poorer cancer survival in Louisiana
Falk et al. (1988)	Life-style risk factors & pancreatic cancer
Fontham et al.(1988a,1993)	Tobacco & cancer
Fontham et al.(1988b)	Dietary vitamins A,C, & lung cancer
Public Data Access, Inc.(1988)	Mortality & toxics along the Mississippi
Olsen et al.(1989)	Mortality experience of a cohort of chemical workers
Chen et al.(1992a)	Tobacco-related cancers
Chen et al.(1992b)	Breast cancer and reproductive system
Fontham et al.(1992)	Gastrointestinal tract cancers
Groves et al. (1992)	Lung cancer & risk factors in St. Bernard
Chen et al. (1994)	An explanation for high cancer death rates
Groves et al.(1996)	Cancer corridor in Louisiana
Chen et al. (1997)	Highlights of cancer incidence in Louisiana
Chen et al. (1998)	Cancer incidence in the industrial corridor (update)
Scroggins JR & Bartley (1999)	Enhancing cancer control: assessing cancer knowledge, attitudes, and beliefs in disadvantaged communities

*Also see Appendix A.

population is identified according to the presence or absence of risk factors associated with the probability of disease occurrence. The conclusions from these studies have not always been consistent. These studies differed in design, size, data source, data quality, length of observation, and statistical procedures, making comparison of results difficult.

Ecologic studies

In ecologic studies on lifestyle, Blot and Fraumeni (1976), using multiple regression analyses, studied the geographic patterns of U.S. lung cancer mortality rates among whites in each U.S. county by demographic and occupational characteristics. The following factors were noted to be related to an increase in lung cancer mortality: urban residence, low educational level, and (among males) the

presence of paper, chemical, petroleum, and transportation manufacturing industries. The largest industry in the 27 Louisiana parishes with the highest lung cancer rates was the chemical industry. Blot and Fraumeni (1978) and Blot et al. (1978) extended this study to bladder and pancreatic cancer. Among males, significant increases in bladder cancer occurred in counties with chemical manufacturing operations. In the analysis of pancreatic cancer mortality rates, no associations were found between industrialization, alcohol intake, or socioeconomic status but they were associated with cigarette smoking.

Voors et al. (1978), using multiple regression analysis, correlated cancer mortality in southern Louisiana parishes with residence in urban areas, residence in wetland areas, and smoking. They found an association in men between respiratory cancer and residence in wetland areas of Louisiana.

Correa and Johnson (1983) investigated respiratory cancer mortality and lifestyle of males in North and South Louisiana (non-Cajuns versus Cajuns) which emphasize the geographic and racial differences in occupation, smoking, diet, and alcohol use. They conducted a survey in four southern and five northern Louisiana parishes and selected a random sample of all males over 30 years of age from holders of driver's licenses with addresses in the parishes selected. The survey indicated that South Louisianians are more frequently drinkers and smokers, start earlier and more frequently use non-filter cigarettes, use more frequently river water as drinking water, and have a higher proportion of workers in oil, shipbuilding, and chemical industry. They concluded that the more industrialized south has higher respiratory cancer rates than the rural north.

Groves et al. (1992) studied the lung cancer rates and risk factors of St. Bernard parish. The parish was divided into six zones, each consisting of one or more census tracts and each containing approximately 10,000 white persons. Occupation, toxic air emissions from industrial plants, and the consumption of diet and tobacco (based on a telephone survey) were investigated as risk factors for lung cancer. Cigarette smoking was specially associated with high lung cancer rates in rural areas.

An ecologic analysis of mortality in counties along the Mississippi River was conducted by Public Data Access, Inc. (1988). Comparisons of mortality rates were made between counties along the Mississippi River and in the U.S. in terms of toxic emissions, toxic waste, pesticide usage, and toxic discharges to the surface waters. Although the report concluded that high overall cancer mortality rates were related to industrial pollution along the Mississippi River, the study has added little new information to the body of knowledge on cancer in Louisiana.

Following this research, studies of cancer incidence in river parishes (East and West Baton Rouge, Iberville, Ascension, St. James, St. John, and St. Charles)(Louisiana State University Medical Center 1993) and in South Louisiana (Groves et al. 1996) were conducted to assess whether incidence data substantiate the reputation of a "cancer corridor" derived from mortality statistics in Louisiana. These studies calculated age-adjusted cancer incidence rates (1983-1987) and compared them with environmental and demographic data and cancer mortality rates. The results were that, with the exception of lung cancer among white men, river parish residents had a lower than average risk of developing the most common types of cancer, and South Louisiana incidence rates were significantly higher than the SEER rates only for lung and larynx

cancers in white males (Groves et al. 1996). In general, these studies concluded the high death rates in South Louisiana are not due to high incidence rates but poor cancer prognosis.

Chen et al. (1998) provided an update of cancer incidence in the industrial corridor for the period 1989-1993. Cancer rates were computed for the entire state of Louisiana and for the combined SEER program. Cancer incidence rates for the industrial corridor were either similar to, or lower than the SEER rates. Only lung cancer in white men and kidney cancer in white women was a significant exception. The findings confirmed the need to develop an impact analysis of environmental exposures and genetic susceptibility on lung cancer risk, and an effective program of tobacco prevention and lung cancer control activities.

Ecologic studies on the relationship between cancer and drinking water in Louisiana have also been carried out. Based on ecologic analyses of essentially the same data, different interpretations and conclusions were stated. These studies were based on regression analyses in which cancer mortality rates by parish in Louisiana for 1950-1969 were used as dependent variables. The percent of the parish population who used the Mississippi River as a source of drinking water, median income, and proportion of rural population, and proportions of the work force engaged in the petroleum, chemical, and mining industries were used as independent variables. Harris (1974) first reported a significant relationship between drinking water and mortality from the urinary tract and gastrointestinal organs among white males.

A more extensive analysis, which included females and nonwhites, was published by Page et al. (1976). This analysis suggested a significantly positive

relationship between water obtained from the Mississippi River and cancer mortality for all sites combined, gastrointestinal organs, and urinary organs. For percentage rural and median income, a relatively significant positive relationship was also found. Except for the chemical industry, occupation did not show any significant effect on cancer mortality. DeRouen and Diem (1977) applied the same multivariate regression model to liver cancer and lung cancer data. The water variable was found to be not significant in any of cancer groups (except for lung cancer mortality in nonwhite females). Furthermore, they compared cancer mortality rates in South Louisiana with those in North Louisiana. Overall, South Louisiana residents had higher cancer mortality rates for kidney, bladder, liver, stomach, and colorectual cancer than residents of North Louisiana. But within South Louisiana, parishes using river water tended to have a higher rectal cancer rate than those not using river water.

In addition to these studies, Chen et al. (1992a, 1992b) and Fontham et al. (1992) reviewed and discussed findings from Louisiana monographs volumes VI and VII of "Cancer Incidence in South Louisiana" (Chen et al. 1990, 1991). They summarized that tobacco-related cancers (for white men), gastric cancer rates (for black men), and pancreatic cancer rates (for whites of both sexes) are significantly higher than those of the nation. Cancers of the breast and reproductive system were less common in South Louisiana than other parts of the nation.

Case-control studies

A series of case-control studies of selected cancer sites in relation to drinking water were conducted by Gottlieb et al. (1981, 1982a, and 1982b). To assess a possible relationship with drinking water source, a comparison of cancer deaths and noncancer

deaths (death certificates) in Louisiana parishes from 1960 to 1975 was carried out. These studies revealed an increased risk for cancer of the rectum in those who consumed chlorinated surface water compared with those who drink groundwater. But during 1983-1987, rectal cancer incidence rates in the New Orleans area paralleled those of the SEER program for males and were lower among females (Groves et al. 1996). The case-control studies provided little definitive information on the relationships between cancer and drinking water in Louisiana.

Case-control studies investigating environment and lifestyle determinants of cancer in Louisiana have focused on three cancer sites: lung, pancreas, and stomach (Chen et al. 1984; Correa et al. 1984a, 1984b, 1985b, 1988; Falk et al. 1988; Fontham et al. 1988a, 1988b, 1992, 1993; Gottlieb and Carr 1981; Groves et al. 1992; Pickle and Gottlieb 1980). These studies all used similar methods. Information on lifestyle exposures was gathered by detailed questionnaires and interviews administered directly to patients or next-of-kin.

Chen et al. (1984), Correa et al. (1984b, 1988), Fontham et al. (1988a, 1988b, 1993), Gottlieb and Carr (1981), and Groves et al. (1992) studied the effects of smoking, occupation, diet, and environmental exposure on the risk of lung cancer in Louisiana. They concluded that smoking is its most important contributor. Nutritional factors (diets low in fresh fruits and vegetables) and occupational exposures (such as to wood dust and mineral oil mist) were found to also increase the risk of lung cancer.

Rothschild and Mulvey (1982) and Mulvey and Rothschild (1983) detected the relationship between occupational histories and lung cancer, through case-control study based on interview data. No association was found between employment in the

petrochemical industry and lung cancer but an elevated lung cancer risk was found among sugarcane farmers who smoke cigarettes.

Studies about lifestyle determinants of pancreatic cancer (Correa et al 1984a, 1988; Falk et al. 1988; Fontham et al. 1988a) did not identify significant association with lifestyle factors. A small but statistically significant elevation in the risk of pancreatic cancer was found among people who are heavy smokers, who have diets rich in pork and other meats, who have low income, and who live in rural areas.

In studies about lifestyle determinants of stomach cancer (Correa et al. 1985b, 1988; Fontham et al. 1992), the results suggested that the causes of stomach cancer are multifactorial and a combination of lifestyle risk factors, such as low intake of fruits and vegetables, excessive salt, high consumption of pork and fat, smoked meat, alcohol, and smoking, influence the development of stomach cancer. But cigarette smoking was not found to be a major factor. According to the studies above, a common finding for cancers of the stomach, pancreas, and lung is that non smoking and fresh fruits (or vegetables), as well as vitamin C and beta-carotene, appear to exert a protective effect.

In case-control studies of death certificates, Shear et al. (1980) found a higher risk of lung cancer associated with residential proximity to cannery industries. Gottlieb and associates (Gottlieb 1980; Gottlieb and Stedman 1979; Gottlieb and Carr 1981; Gottlieb et al. 1982c) noted that an increased risk of lung cancer is associated with not only residential proximity to petroleum and chemical industries but also employment in the petroleum, shipbuilding, and fishing industries. But these studies did not provide proper analysis of occupation and residence because the data were based only on death certificates (Wong and Foliart 1989).

Occupational cohort studies

Several historical cohort mortality studies have been conducted in Louisiana. These studies include a variety of industries, such as petroleum products, petrochemicals, agricultural chemicals, insecticides, synthetic rubber, and isopropyl alcohol. Except for the Exxon refinery/chemical plant study (Hanis et al. 1982), most of the studies were relatively small. Lynch et al. (1979) found a significant excess of laryngeal cancer among Exxon employees at the Baton Rouge isopropyl alcohol plant. Weill et al. (1979) and Hughes et al. (1987) studied the cancer mortality of employees of two New Orleans asbestos cement products manufacturing plants. A significant lung cancer risk was found among employees at the plant that uses both crocidolite and chrysotile asbestos fibers. In other refinery/chemical plant studies (Herman 1981; Hanis et al. 1982; Holmes et al. 1986; Ference et al. 1987; Olsen et al. 1989), there was no observed mortality excess from either cancers of all sites combined or lung cancer. Hanis et al. (1982) noted that significant excesses of kidney and pancreatic cancers were detected in the Exxon Baton Rouge refinery.

2.4 Summary

This chapter reviewed existing literature for the study of cancer mortality and environment. In summary, human, disease, and environment are closely related to one another. The major approach to studying the relationships of human, disease, and environment involves the use of spatial analysis in analyzing disease patterns. Spatial analysis techniques include disease mapping and associative analysis. In particular, the application of GIS makes it possible to analyze complicated spatial relationships effectively between disease and environmental factors.

Cancer is the leading cause of death in the U.S. (and most other nations) as well as in Louisiana. A number of statistical or descriptive studies have been conducted in Louisiana. The conclusions from these studies have not always been consistent. These studies differed in design, size, data source, data quality, length of observation, and statistical procedures. But most studies have indicated that southern Louisianians have had higher cancer mortality rates than other U.S. residents. If the cause and nature of certain cancers attributed to environmental factors were known, preventative guidelines might be offered for many cancers.

CHAPTER 3

MATERIALS AND METHODS

The previous chapters provided the objectives and hypotheses of this research and outline fundamental literature dealing with the issues of cancer and environment. This chapter contains detailed descriptions of the study area, data sources, methods, and limitations of the present study.

3.1 Study Area

To provide a better understanding of how cancer mortality rates and their spatial distributions in Louisiana compare with those in the nation and to identify the association between cancer mortality rates and environmental factors in Louisiana, geographical study areas used in this research were considered at different levels: i.e. state, county (parish), and census tract.

State level: States included were the 48 contiguous states and the District of Columbia (Figure 3.1). The District of Columbia was considered a 'state' for purposes of this study. The state geocode consists of a 2-digit Federal Information Processing Standard (FIPS) code.

County (parish) level: The first order division of each state contains counties for the states (parishes for Louisiana). This study covered 3,111 U.S. counties from 1953 to 1987 (Figure 3.1) and 64 Louisiana parishes from 1953 to 1993 (Figure 3.2). FIPS county codes are 3-digit numbers assigned within states.

As designated by Louisiana Tumor Registry (Chen et al. 1996), South Louisiana includes 35 of 64 parishes. It covers five regions: New Orleans (Jefferson, Orleans, and St. Bernard); Baton Rouge (Ascension, Assumption, East Baton Rouge, East Feliciana,

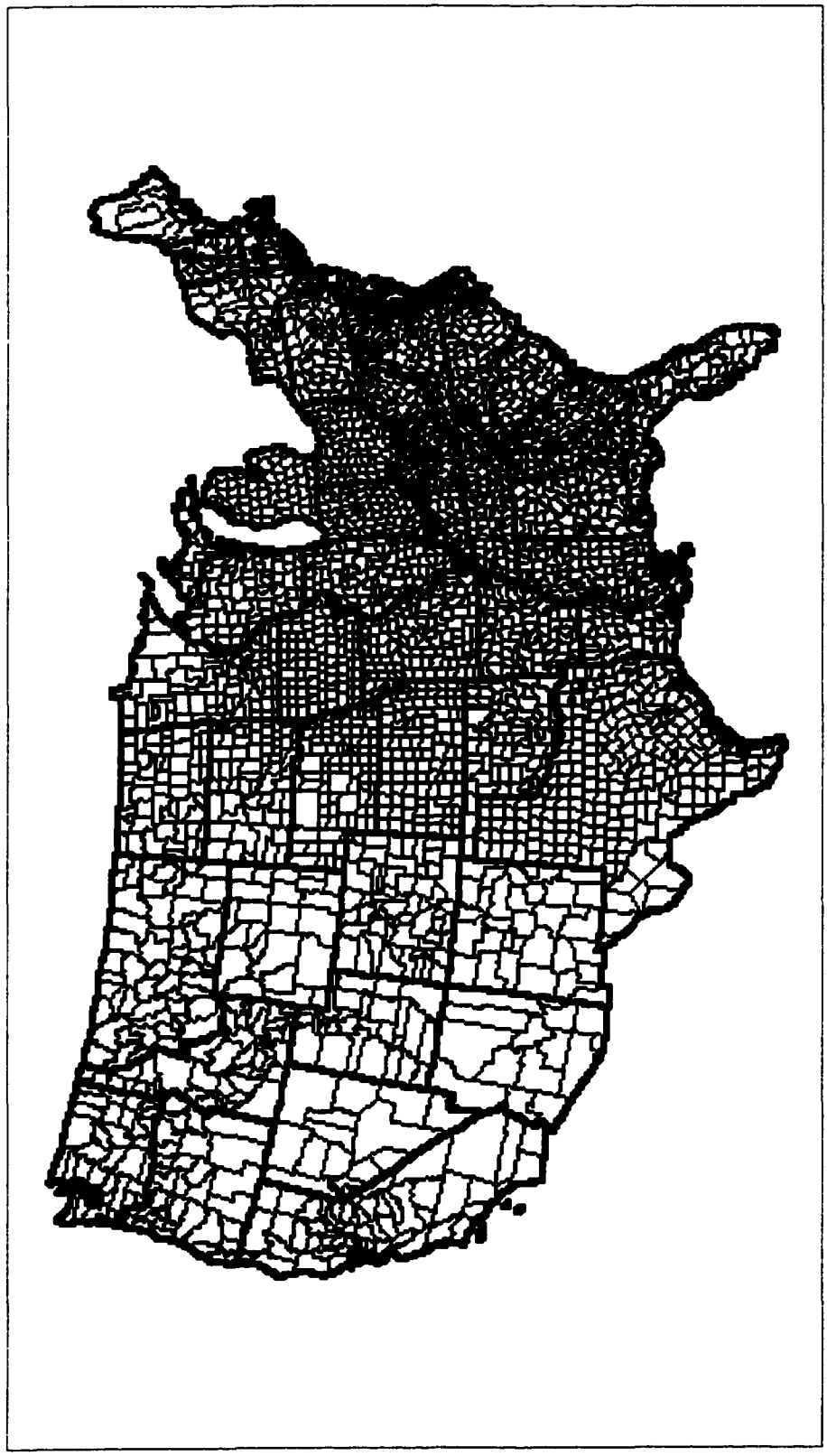

Figure 3.1 U.S. State and County Boundaries
(Source: Modified by Author from U.S. Census Bureau, 1990)

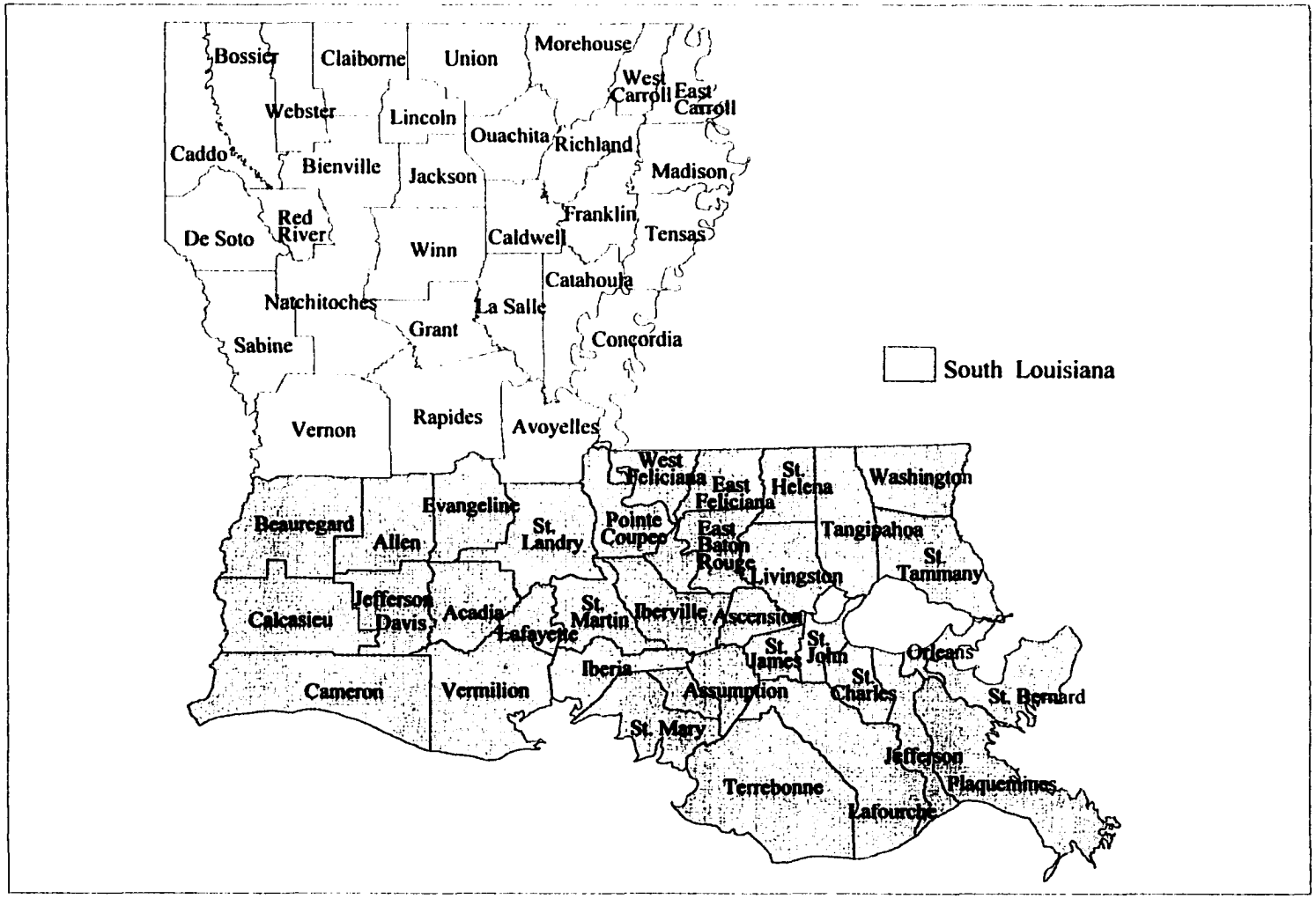

Figure 3.2 Louisiana Parish Boundaries and Names (Bark, 1999)

Iberville, Livingston, Pointe Coupee, St. Helena, Tangipahoa, West Baton Rouge, and West Feliciana); Southeast Louisiana (Lafourche, Plaquemines, St. Charles, St. James, St. John, St. Tammany, Terrebonne, and Washington); Acadiana (Acadia, Evangeline, Iberia, Lafayette, St. Landry, St. Martin, St. Mary, and Vermillion); and Southwest Louisiana (Allen, Beauregard, Calcasieu, Cameron, Jefferson Davis).

Census tract level: Census tracts are small, relatively permanent geographic subdivisions of a county or equivalent entity. The U.S. Census Bureau has defined that tracts are compact contiguous areas with populations of about 4,000 persons and that the area should, if possible, try to avoid combining non-homogeneous areas. The ideal census tract would be a locally recognized "neighborhood" within a city. Census tracts are assigned 4-digit numeric codes. Census tracts can also have a 2-digit suffix code, usually indicating a 'split' of a tract from an earlier census year. The 1,102 Louisiana census tracts (from 1988 to 1992) appearing in this study are those defined at the time of the 1990 census (Figure 3.3).

3.2 Data

This research incorporated several sources of data collected on multiple scales, for different periods, from several agencies and documents (Appendix B). Data on cancer mortality and incidence rates (from 1953 through 1993), as well as environmental variables (from 1980 through 1989) in Louisiana, and cancer mortality rates in the U.S. (from 1953 through 1987) were assembled.

Data in limited spatial scales and temporal spans were used for the study because data are missing for a certain period and are not consistent and accurate. Although there have been a constant number of counties, some name and administrative status changes

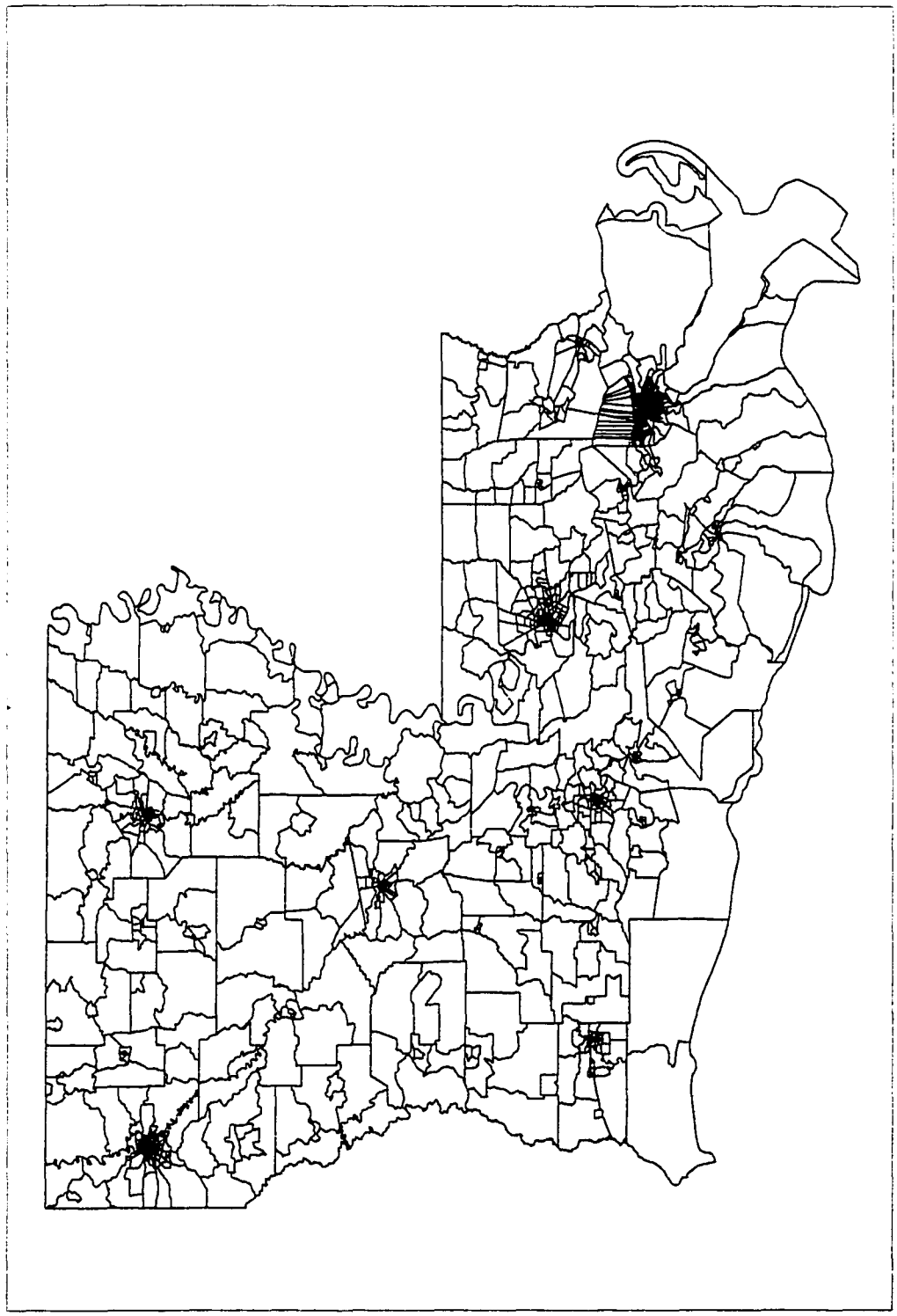

Figure 3.3 Louisiana Tract Boundaries
(Source: Modified by Author from U.S. Census Bureau, 1990)

have occurred over the years. All of these changes have been removed or adjusted, and the data sets have been shown in Figures 3.1-3.3. The collection, processing, and calculation of data are explained below in detail.

3.2.1 Cancer Data

Cancer mortality rates are often used because they are available for all states and are easily accessible. The death rate measure not only the underlying risk of getting the disease but also of survival. Cancer incidence, on the other hand, refers to diagnosed cancer cases and would be a better indicator of the risk of developing cancer in a population.

In this study, all cancer mortality (incidence) rates were the number of cancer deaths (occurrences) per 100,000 population and age-adjusted to the 1970 U.S. standard population. Age adjustment is a standard statistical practice to allow comparison of population with different age distributions. Of all cancer sites, five most common cancers in the U.S. and Louisiana were studied, including:

Lung: Lung cancer is a major cause of death in the U.S. as well as Louisiana. It is one of the most preventable of all malignancies. Lung cancer in this study included cancer of the trachea and pleura, as well as the lung.

Colorectum: The colon and rectum are often lumped together as "colorectal cancers" and make up the large bowel, or large intestine. The colon refers to the upper five or six feet of the large intestine, and the rectum to the last five or six inches.

Breast: Breast cancer is one of the most common cancers among U.S. and Louisiana women.

Prostate: Cancer of the prostate is one of the most common cancers among U.S. men. The prostate gland, located at the base of the penis, surrounding the urethra, produces seminal fluid.

Stomach: Before World War II, stomach cancer was the leading cause of cancer death in the U.S. Since then, there has been a dramatic decrease in stomach cancer death rates for both sexes. Currently, the U.S. death rates for stomach cancer are among the lowest in the world, but Louisiana has had higher stomach cancer rates than the nation.

Cancer Mortality Data

1. U.S.: The age-adjusted cancer mortality rates by state and county for 1953 through 1987 were obtained from National Technical Information Service (NTIS). These data were provided in 5-year periods; for example, from 1953 through 1957, 1958 through 1962, up to 1983 through 1987[3.1]. For comparison with Louisiana data, the age-adjusted cancer mortality rates by site, race (whites and nonwhites only), and sex were grouped into three periods, 1953 through 1977, 1978 through 1987, and 1953 through 1987.

2. Louisiana parishes: To correspond to national data rates, Louisiana cancer rates by parish were grouped into three periods, 1953 through 1977, 1978 through 1987, and 1953 through 1987 from the cancer data set of NTIS. Age-adjusted cancer mortality rates by parish were acquired from the Centers for Disease Control and Prevention's

[3.1] Also, cancer mortality rates for all the U.S. counties including Louisiana for the period 1950-1979 were available from the U.S. Cancer Mortality Rates and Trends, 1950-1979 (Riggan et al. 1983).

(CDC) Wonder system to allow more consistent and accurate analysis with Louisiana environmental data for the 1980s and to compare with cancer data of Louisiana census tract from 1988 through 1993. CDC Wonder, which may be queried online (*http://wonder.cdc.gov/wonder/data/mortJ.shtml*), is an easy-to-use system that provides a single point of access to a wide variety of CDC reports, guidelines, and public health data.

Also, raw data for cancer mortality of all cancer sites (International Classification of Disease (ICD) -9 code; 140-208) by Louisiana parish from 1950 through 1993 were obtained from the Louisiana Department of Health and Hospitals (LDHH), Office of Public Health, Division of Public Health Statistics. Data were in single record format as an American Standard Code for Information Interchange (ASCII) file that contained the year of death, age, race, sex, parish of usual residence, and specific cause of death for each individual who died. Personal information such as name, social security, and street address, was not included.

To allow more accurate analysis and comparison with Louisiana cancer data, age-adjusted cancer mortality rates by site, race (whites and nonwhites only), and sex in 64 parishes from 1980 through 1989 were calculated using Statistical Analysis System (SAS) (Version 6.10) (Appendix C). Rates were computed for five important types of cancer as well as for all cancers combined. Intercensal population estimates by 5-year age groups for each year from 1980 through 1989 were obtained from the Census Bureau via the National Cancer Institute (NCI). The statistical methodology used to calculate the cancer mortality rates was the same as those used in previous NTIS reports.

The average annual age-adjusted (1970 U.S. standard population) mortality rates (per 100,000) were calculated using the following formulas.

Methods (Calculation of age-adjusted rates (direct method)) (NTIS 1992)

(3.1) $$S_{i,t} = \frac{d_{i,t}}{p_{i,t}}$$

(3.2) $$A_t = \frac{\sum_{i=1}^{n} (S_{i,t} * P_i)}{\sum_{i=1}^{n} P_i} * 100{,}000$$

(3.3) $$A_d = \frac{\sum_{t=1}^{m} A_t}{m}$$

Where $S_{i,t}$ = the age-specific death rate in age group i over t years;
A_t = the age-adjusted death rate per 100,000 over t years;
A_d = average annual age-adjusted death rate.
$d_{i,t}$ = the number dead in age group i over t years;
$p_{i,t}$ = the number in age group i over t years;
P_i = the number in age group i in the standard population;
t = the number of years observed;
i = a particular age group 5 years wide;
n = the number of age groups;
m = (10), number of years in period

3. Louisiana census tract: Raw data for all cancer sites by tract in Louisiana from 1988 to 1993 were obtained from LDHH as an ASCII file. Data were in single record format and contained the year of death, age, race, sex, tract of residence, and specific cause of death for each individual who died. Eighty-two percent of the data included information on the tract of usual residence.

Cancer Incidence Data

For further comparison, age-adjusted (U.S. 1970 standard population) cancer incidence data by parish for the period 1983-1987 were obtained from the Louisiana Tumor Registry (LTR), Office of Public Health. However, such cancer incidence data were available only for South Louisiana. Average annual age-adjusted cancer incidence by region, including all of Louisiana (by region) from 1988 to 1992, was obtained from the report *Cancer Incidence in Louisiana, 1988-1992* (Chen et al. 1996).

LTR is a population-based cancer registry which was established in 1974 to register cancer cases in the Orleans parish area, expanded to South Louisiana in 1983 and to statewide coverage in 1988.

3.2.2. Environmental Data

Environmental data were collected for all parishes in Louisiana from 1950 to 1993, but this study used and explained environmental data by Louisiana parish in only the 1980s (for multiple regression analysis), to identify the relationship of cancer and the environment. The primary reason for using data at the parish level is data availability, as consistent data in spatial and temporal scales are not yet available at the smaller than parish level.

Even though the examination of data at the parish level leaves crucial questions unanswered, it could provide useful information that exists across broad geographic areas of the state. When significant relationship or variation between cancer and potential environment factors is discovered across parishes, it indicates definitely the need to examine such variation at smaller levels of aggregation such as the census tract and census block levels.

Several types of data considered at the parish level were as follows:

1. Population: Data of intercensal estimated population (released by U.S. Census Bureau) from 1980 to 1993 were provided by NCI and the Division of Business Research, Louisiana Tech University. Population includes age group (5-year intervals), sex, and race. Data for 1981-1989 are intercensal estimates of the July 1 resident population. The 1980 and 1990 estimates are April 1 modified census counts. The estimates for 1991-1993 are postcensal estimates of the July 1 resident population. The number of population from 1980 to 1989 was averaged for the 1980s.

The U.S. 1990 population obtained from the U.S. Census Bureau was used for demographic data by census tract from 1988 to 1993. Also, the percentages of white and nonwhite were collected from the *County and City Data Book* (U.S. Census Bureau 1984, 1988, and 1994) and averaged for the 1980s.

2. Land area: Square miles for each parish were acquired from the USA Counties on CD-ROM databases (U.S. Census Bureau 1994) for 1980 and 1990. These land areas were averaged to represent the 1980s.

3. Population density: The data for population density (population per square mile) were acquired from the USA Counties on CD-ROM for 1980 and 1990. The density figures were averaged to represent the 1980s.

4. Income: Per capita income is defined as real income divided by total population. Per capita income values from 1980 to 1990 were obtained from the Regional Economic Information System, Bureau of Economic Analysis, U.S. Department of Commerce (1994 May), through the Division of Business Research, Louisiana Tech University. They were averaged for the 1980s.

5. Education status: Percentages of high school graduates or higher education attainment were collected from the USA Counties on CD-ROM for 1980 and 1990. These percentages were averaged to represent the 1980s.

6. Poverty (Persons below poverty level): The percentages of persons with income below poverty level in 1979 and 1989 were acquired from the *County and City Data Book* (U.S. Census Bureau 1984 and 1994). These percentages were averaged to represent the1980s.

7. Occupation: The numbers of persons employed by important industry were taken from the Regional Economic Information Systems (REIS) on CD-ROM, *Louisiana Employment and Wages Annual Report, County and Business Patterns* for 1980 to 1989, and *the Census of Population* for 1980 and 1990. To keep a consistent number of employees and same classifications of industry, the numbers of persons by occupation were used from *the Census of Population*, and the numbers of employees from other sources were compared for the accuracy and missing value of data. To adjust for the variation in size among parishes, the number of employees was divided by the population of the parish in the 1980s, and the result was multiplied by 100. The occupational classifications covered in this research are as follow:

• Agriculture: Farming activity such as cultivating the soil, producing crops, and raising livestock.

• Construction: General contractors in the construction of buildings, highways, bridges, and special trades such as plumbing and painting.

• Manufacturing: Production of a finished product which is ready for utilization or consumption, and a semi-finished product which is transformed from a raw material.

Manufacturing in this study was divided into manufacturing (including all products) and chemical manufacturing (including only chemical and allied products).

• Mining: Extraction of minerals such as copper ores, lignite mining, crude oil and natural gas production and natural gas liquids, oil and gas field contract field services, salt, sulfur, sand and gravel quarries.

• Transportation: Air, water, truck, bus and pipe lines with related services.

• Health services: Hospitals and health services.

• Education services: Elementary and secondary schools, colleges, and other educational services.

8. Toxics Release Inventory (TRI): Toxic chemical pollutants are generated from many different sources, including manufacturing and non-manufacturing industrial processes, use and disposal of consumer products, use of agricultural chemicals, and mobile sources such as auto mobiles. Congress passed the Emergency Planning and Community Right to Know Act (EPCRA) in 1986 as Title III of the Superfund Amendments and Reauthorization Act (SARA). This Act requires manufacturing facilities that manufacture, use, or process over 25,000 pounds of any reportable toxic chemicals per year to submit an annual report estimating the quantities of toxics released into the air, water, and land environment, or transported offsite. Air emissions include leaks, evaporative losses, and stack emissions. Land emissions cover surface impoundments, landfills, and land treatment. Water emissions are discharges into any surface water body, including storm water runoff that is contaminated with the listed chemicals (Louisiana Department of Environmental Quality (LDEQ) 1991).

It should be recognized that TRI data have some limitations. They include only manufacturing facilities, and thus exclude other important sources of pollution. The data-collection system relies on self-reporting and permits estimation. One hundred eighty-five different substances were included in the list of chemicals reported in Louisiana from 1987 to 1989, and all measurements were reported in pounds per year. Use of such simple aggregates as pound measurement may provide misinformation about exposure to environmental hazards. For example, a large volume of a less toxic substance may mistakenly be perceived as more serious than a smaller release of more highly toxic substance. However, the TRI data are the most comprehensive data currently available on potential environmental exposure to toxic chemical releases.

For this study, TRI data from 1987 to 1989 were obtained from the Technical Program Support Division of the Louisiana Department of Environmental Quality. The database contains a record for each substance released by each facility. Each facility's street address, zipcode, and parish code were reported. There were two measures of potential environmental exposure for each medium (releases by air, water, underground injection, and land) and for the total of the four media. The first was the total (in pounds) of all releases to that medium within each parish. The second was the total, by parish and by medium, of all releases of chemicals classified as human carcinogens and probable human carcinogens according to the EPA (See Table 2.3 in Chapter 2). To normalize the variation in size among parishes, the number of pounds released was divided by the land area as the 1980s.

9. Waste Sites: Louisiana's environmental waste sites can be divided into 3 categories - solid waste, inactive and abandoned waste, and hazardous waste sites.

Solid waste includes solids, liquids, and gases which have not been classified as hazardous. Some solid waste is very toxic. A study found toxic and cancer-causing chemicals in the wastewater of all the solid waste landfills may cause as great a cancer risk as those from hazardous waste landfills (Brown and Donnelly 1988). Solid waste is disposed of in landfills, open dumps, incinerators, and landfarms. The total number of sites in Louisiana (in the 1980s) was 1,752. Eighty percent of all solid waste disposal facilities in the state were either landfills or surface impoundments. Landfills are by far the most widely used disposal option. The number of landfills in Louisiana, as in the rest of the nation, is diminishing. In 1981, there were some 788 landfill sites across the state (Clipp 1994). In 1990, there were fewer than 30. The lists of solid waste sites in Louisiana were obtained from the Solid Waste Division of Louisiana DEQ. As a measure of potential exposure, the number of solid waste sites in each parish was divided by the land area of the parish, to give the number of sites per square mile. The result was multiplied by 1,000 in order to avoid very small numbers.

Inactive and abandoned sites are areas where hazardous waste was disposed of or handled improperly. The lists of inactive and abandoned hazardous waste sites in Louisiana were obtained from Inactive and Abandoned Hazardous Sites Division of Louisiana DEQ. The sites used in this study were a subset of the Comprehensive Environmental Response, Compensation and Liability Act Information Systems (CERCLIS) -listed sites (a list of potentially hazardous waste sites throughout the nation) and included 506 sites in Louisiana for the 1980s. The number of inactive and abandoned hazardous waste sites in each parish was divided by the land area of the parish and the result was multiplied by 1,000 in order to avoid very small numbers.

Hazardous waste is the type of waste that is normally considered dangerous and is typically manufactured and disposed of by industry. For example, benzene, used in manufacturing a variety of products, is a known cancer-causing chemical and is listed as a hazardous waste. Once on the CERCLIS list, sites are investigated by the EPA according to the "hazardous ranking system" to evaluate and rank the sites based on the seriousness of human exposure to toxics contained onsite. The only outcome of the ranking available is whether or not the site is placed on the National Priorities List (NPL) and thus becomes eligible for remediation under the superfund. To be eligible for the superfund listing, a site must score above 28.5 in the EPA's Hazard Ranking System. Louisiana had 6 superfund NPL sites and 5 proposed NPL sites in seven different parishes during the 1980s (Table 3.1).

Data on the geographic locations of some existing superfund sites were not clear from the Hazardous Waste Division of LDEQ in 1997. Thus, these sites were redigitized by the latitude and longitude information on maps obtained from the Hazardous Waste Division of Louisiana DEQ. The number of parishes included in the distance (a cirlce) of 2 miles from the centroid of each superfund site was calculated by a GIS software (Windows NT Intergraph) (See Appendix D for the procedure and commands). As another indicator of potential environmental exposure, the number of such hazardous waste sites in each parish was divided by the land area of the parish. The result was multiplied by 1,000 in order to avoid very small numbers.

10. Pesticides: Pesticide refers to any substance used to kill unwanted organisms. It includes chemicals produced specifically to kill fungi, insects, weeds, rodents, mites, or bacteria. Currently, over 1 billion pounds of pesticides are used annually in the U.S.

Table 3.1 Louisiana's Superfund Sites

Superfund National Priorities List (NPL) Sites

Site	Latitude, Longitude	Parish	
Bayou Bonfouca	30:16:56, 89:47:06	St. Tammany	Abandoned wood treating facility
Bayou Sorrel	30:13:24, 91:24:42	Iberville	Hydrocarbon processing waste disposal site
Cleve Reber	30:09:22, 90:52:17	Ascension	Industrial landfill for petrochemical waste
Dutchtown Treatment	30:15:21, 90:58:48	Ascension	Oil reclamation facility
Old Inger	30:09:55, 90:59:42	Ascension	Refinery which used other refineries' waste disposal site
Petro-Processors, Inc.	30:34:54, 91:14:37; 30:35:20, 91:13:08	East Baton Rouge	Petrochemical waste disposal site

Proposed NPL sites

Site	Latitude, Longitude	Parish	
Combustion, Inc.	30:30:41, 90:53:51	Livingston	Waste oil refinery
DL Mud	29:57:17, 92:11:09	Vermilion	Drilling mudmixing facility
Gulf Coast Vacuum Services	29:57:29, 92:11:03	Vermilion	Waste oil handling facility, drilling fluid waste disposal facility, and truck washout facility
Louisiana Army Ammunition Plant	32:33:21, 93:25:49	Webster	Wastes from munitions manufacture, a burning ground, a landfill, lagoons, and an oily waste landfarm
Pab Oil	30:00:57, 92:06:47	Vermilion	Oilfield waste disposal facility

(Source: Modified from Hazardous Waste Division of Louisiana DEQ, 1997)

Of these, 59% are herbicides, 24% are insecticides, and 16% are fungicides or other biocides. The vast majority (80%) are used in agriculture, but a substantial amount are also used by government, industry, and homeowners. Although most pesticides are used for growing crops, the concentration of pesticide applications in urban areas is often much higher than in rural areas. Pesticides are used in urban areas for care of homes, lawns, and gardens, as well as for ground skeeping, mosquito control, and structural pest control. Herbicides are also sprayed on roadsides, utility and railroad rights of way, and golf courses. There are about 700 pesticides in use in about 40,000 products. Of the 700, only a few dozen have been "reregistered," or evaluated by the U.S. EPA for adverse health effects. Even a pesticide that passes this evaluation is not necessarily a safe product (Clipp 1994).

The use of pesticides in the U.S. has grown astronomically since their commercial introduction prior to World War II. Despite an almost tenfold increase in pesticide use, crop losses from insect pests doubled between the 1940s and 1970s. The Louisiana Department of Agriculture and Forestry (LDAF) oversees the use and registration of pesticides in Louisiana.

The pesticides referred to in this study consisted of agricultural chemicals only (including fertilizer and lime) while TRI data included the chemicals from manufacturing facilities only. The pesticide data were derived from Census of Agriculture, 1982 and 1987 because about 80% of all pesticides were used in agriculture and no accurate data for all pesticides existed in the 1980s. County pesticide usage was measured by the aggregate number of acres where agricultural chemicals were used. To normalize the variation in size among parishes, the number of averaged acres was

divided by the land area of the parish in the 1980s, and the result was multiplied by 1,000.

11. Wetlands: Coastal Louisiana contains 40 % of the U.S. wetlands and nearly 4 million acres of saltwater, brackish, and freshwater wetlands. Wetlands are lands where saturation with water is the dominant factor determining the nature of soil development and the types of plant and animal communities living in the soil and on its surface. Exact definition and classification of wetlands are difficult because a wetland is a complex ecosystem with intricate relationships of water, soil, and vegetation. The term 'wetland' in this study followed one of the classifications by Cowardin et al. (1992): (1) areas with hydrophytes and hydric soils, such as those commonly known as marshes, swamps, and bogs; (2) areas without hydrophytes but with hydric soils; (3) areas with hydrophytes but nonhydric soils; (4) areas without soils but with hydrophytes such as the seaweed-covered portions of rocky shores; and (5) wetlands without soil and without hydrophytes, such as gravel beaches or rocky shores without vegetation. GIS software (Windows NT Intergraph, Erdas Imagine, and Unix ArcInfo) were used to calculate the area of wetlands of each parish from the classified habitat file for 1978 and 1988 from the Louisiana Department of Natural Resource. The area of wetlands in each parish was divided by the land area of the parish, and the result was multiplied by 100. More detailed processes are listed as Appendix E.

12. Urban population: The term 'urban' means the characteristic of being situated in a city or town. Land use in a densely populated area or place containing 2,500 or more people is considered to be urban in nature and its population urbanized. Any sparsely settled area or place containing fewer than 2,500 people is designated as

rural. Rural land use and population are subdivided into rural farm and non-farm (U.S. Census Bureau 1995). The urban population (1980 and 1990) was derived from *Louisiana FactBook* (1993). The parish population living in urban area was averaged. The result was divided by the land area of the parish in the 1980s and then multiplied by 100.

13. Air Quality (Ozone): Concern about air pollution exists because polluted air can damage the health of humans and the environment. Airborne toxics, such as Toxic Release Inventory (TRI) chemicals, pose a serious pollution control problem because toxic chemicals are emitted into the atmosphere by many activities (*Louisiana Toxics Release Inventory* 1993). The National Ambient Air Quality Standards (NAAQS) set by the EPA lists six air pollutants (ozone, sulfur dioxide, nitrogen oxide, carbon monoxide, total suspended particulates, and lead) as posing the greatest overall threat to air quality (U.S. Environmental Protection Agency 1988). LDEQ maintains a network of air quality monitoring stations which measure the levels of criteria pollutants in the air on a daily basis. Monitors are generally located in areas where the potential risks to human health from exposure to air pollution are the greatest.

The pollutant measured at the most stations, and the only one that exceeded federal ambient air quality standards at any station, was ozone. High levels of ozone can produce respiratory lung hazards such as shortness of breath, wheezing, congestion, sore throats, and burning eyes. Ozone is a particular problem for asthmatics and those prone to allergies. The number of days in the year during which monitoring stations within the parish reported values in excess of the federal standard of 0.12 parts per million (ppm) was used as a measure of potential exposure. The number of ozone exceedance days

from 1980 to 1990 was taken from the homepage of Air Quality Division of DEQ (www.deq.state.la.us/oarp/03exceed/html). If the parish contained more than one monitoring station, the number of days of ozone exceedances at each station were averaged.

14. Water Quality (Drinking water): Louisiana has abundant water resources throughout the state. Water use is defined as water withdrawn or diverted from a ground- or surface-water source to be used for public supply, industry, power generation, rural domestic, commercial, industrial, and public water use. Five-year investigations have been published in a series titled, "*Pumpage of Water in Louisiana*" or "*Water Use in Louisiana* "(Neil et al. 1991). Eighty-six percent of the total Louisiana population is supplied with water from a public supplier. Public supply refers to water withdrawn and delivered to a group of users by public and private water suppliers. Of the public water users, 55% were supplied with water from a ground-water source and 45% were supplied with water from a surface-water source. From 1985 to 1990, ground-water withdrawals in Louisiana decreased by 6.8%, and surface-water withdrawals decreased by 10%. Total water withdrawals in Louisiana decreased by 10% from 1985 to 1990. Forty-five percent (610 million gallons per day (Mgal/d)) of all ground water withdrawn was from the Chicot aquifer system. Another 21 percent (280 Mgal/d) was withdrawn from the Mississippi River alluvial aquifer. Seventy-two percent (5,800 Mgal/d) of all surface water withdrawn was from the Mississippi River (Neil et al. 1991). The Mississippi River provides the largest source of water for public supplies. Approximately 3,100 square miles of Louisiana is covered by surface water.

Water pollution by chemical carcinogens results in human health risks not only by exposure through drinking water from surface or ground water supplies but also through organisms in the food chain which may be contaminated with carcinogens. Drinking water in Louisiana contains complex mixtures of known and suspected carcinogens including asbestos, metals, radioactive substances, and industrial chemicals. Generally, surface water contamination has been higher than ground water contamination. More river water is used for drinking purposes in the south and more private wells in the north. Many of these people depend upon the Mississippi River as a source of drinking water.

ASCII delimited files of Louisiana's community nitrate sampling data from the late 1970s through the late 1980s were obtained from the Safe Drinking Water Program of Louisiana Office of Public Health. Only Louisiana's community surface water data were used, because metal and nitrate sampling data for drinking water did not contain radiological data. The number of people who drink from surface water systems in each parish was divided by its 1980s', and the result was multiplied by 100.

15. Geographic Location: Area based information may be aggregated and represented by one single geographical location. The geographic coordinates can be specified in latitude and longitude. For the geographic location of each census area (parish and tract) in Louisiana, the latitude and longitude of the area centroid were extracted from the *Mable/Geocorr V2.5-Geographic Correspondence Engine* (*http://www.census.gov/plue*). Land area -weighted method was selected in weighting the centroids.

3.3 Methods

Spatial analyses for the patterns and relationships of cancer mortality and environment included statistical mapping and associative analysis. The statistical analyses of this research were done using GIS softwares (Intergraph, ArcInfo, and Arcview) and statistical software (Statistical Package for the Social Sciences (SPSS), Statistical Analysis System (SAS), and SaTScan). Data were analyzed according to specific purpose of each part of the study. For Hypothesis One, statistical mapping, significance test of rate differences, and factor analysis were used to describe the spatial distribution of higher cancer mortality rates and trends in Louisiana compared with those in the nation. Hypothesis Two about the clusterings of cancer (breast, colon and rectum, lung, prostate, and stomach) mortality rates was tested using factor analysis, spatial autocorrelation analysis, and scan statistic analysis. For Hypothesis Three, which assumed that cancer mortality patterns were associated with environmental variables, factor analysis and multiple regression were used.

3.3.1 Statistical Mapping for Patterns of Cancer and Environment

Statistical maps of mortality and environmental factors are useful aids in describing the distributions, relationships, and trends in health, disease, and environment. There are several classification methods used in statistical mapping based on the nature of data and purpose. The classification method determines the appearance and message of a map and affects the visual interpretation. It is illustrated below how the same set of attribute values are divided into classes by different classification methods.

The equal area method classifies polygon features by finding breakpoints in the attribute values so that the total area of the polygons in each class is approximately the same. This tends to hide the variation of values between smaller areas. The quantile classification method assigns the same number of features in each class. The classes can, therefore, be misleading because low values are often included in the same class as high values. The equal interval method divides the range of attribute values into equal sized subranges. It is not good to reveal subtle differences between features with similar values. The natural breaks method identifies breakpoints (by statistical formula of Jenk's optimization that minimizes the variation within each class) by looking for groupings and patterns inherent in the data (Environmental Systems Research Institute, Inc. 1996). Poisson probability distribution is used to calculate the probability of any given value level. It describes the likelihood of the occurrence of rare, random events, given a mean expectation and a variance. The standard deviation method uses the standard deviation of the distribution of rates as a basis for calculating standard distances in terms of the standard normal curve (Armstrong 1969). The selection of a probability method for a specific problem would depend upon, firstly, the normality of values and, secondly, the consistency and comparability of the size of numbers in the observed and expected columns.

As an important classification method that enables cross comparison, the standard deviation method was used in this study to establish class intervals for mapping. The standard deviation method permits more valid comparison between maps of different rate dispersion on different subjects. But a measure of significance based upon deviations from the mean is not entirely satisfactory when the mortality rate values

are not normally distributed. In general, the standard deviation as a measure of significance is more likely to suit common complaints and diseases of high mortality (McGlashan 1972). Therefore, there would seem to be an advantage in basing class intervals on a measure of the dispersion of a particular set of rates, rather than on percentiles, or arithmetic divisions of the measurement scale, which may not be related to the dispersion characteristics of distribution. All cancer mortality rates and environmental variables for this dissertation were mapped using a standard deviation classification by Arcview (Release 3.0). If the variables showed little indication of normal distribution, then the quantile method was used for mapping the cancer mortality rates and environmental factors.

Tests of Normality

Statistical mapping has different classification methods according to the nature of data. The normal distribution is very important to statistical inference. To examine the assumption that data represent a normal distribution, a normal probability plot is used. Although normal probability plots provide a visual basis for testing normality, it is desirable to compute a statistical test of the hypothesis that the data are from a normal distribution. Two commonly used tests are the Shapiro-Wilk test and the Lilliefors test. A Shapiro-Wilk statistic is calculated for samples with 50 or fewer observation whereas a Lilliefors test is used for samples of more than 50 observations. The Shapiro-Wilk test shows good power in many situations compared with other tests of normality (Norusis 1992). The Lilliefors test is a modification of the Kolmogorov-Smirnov test and is used when means and variances are not known but must be estimated from the

data. The Kolmogorov-Smirnov test is based on the largest absolute difference between the observed and the expected cumulative distributions.

To examine the normality of cancer data and environmental variables for this study, normality plots and the Lilliefors (or Shapiro-Wilk) test were used. The Kolmogorov-Smirnov statistic, with a Lilliefors significance level, is displayed in Appendix F. Tests of normality for the U.S. cancer rates by state (and county) and Louisiana cancer rates by parish from 1953 to 1987 showed normality. Thus, these rates were mapped by the standard deviation classification method.

The hypothesis of normality is rejected when significance levels are small (i.e., $p < 0.01$). However, it was almost impossible to find data that are exactly normally distributed because almost any goodness-of-fit test results in rejection of the null hypothesis when the sample size is large. Therefore, for large data sets, you should look not only at the observed significance level but also at the actual departure from normality. And it is sufficient that the data are approximately normally distributed for most statistical tests (Norusis 1992). In this study, most of the data (except for drinking water, population density, TRI, wetlands, etc.) are approximately normally distributed.

3.3.2 Factor Analysis

Factor analysis is a common statistical technique to analyze interrelationships among many variables in psychology and economics as well as in medicine. This procedure identifies underlying factors that explain the correlations among a set of variables. Its purpose is often to summarize a large number of variables using a smaller number of factors. Factor analysis can also function to eliminate or reduce multicolinearity in a data set, through the transformation of variables into several factors

based upon their covariances. The basic assumption of factor analysis is that the underlying dimensions, or factors, can be used to explain complex phenomena.

In this research, factor analysis was used to determine whether the geographical distributions of cancer mortality rates of various sites and environmental variables could be explained by a few common factors. The SPSS statistical package (Release 8.0) was used to carry out the factor analysis in this study.

The mathematical model for factor analysis is similar to a multiple regression model. The model assumes that the original variable is influenced by various determinants: a part shared by other variables, known as the common variance, and a unique variance that is the residual from the multiple relationships, and also that relating to measurement error.

(3.4) $$X_i = A_{i1}F_1 + A_{i2}F_2 + \ldots + A_{ik}F_K + U_i$$

The F's are the common factors, the U is the unique factor, and the A's are the constants used to combine the k factors.

(3.5) $$F_j = \sum_{i=1}^{p} W_{ji} X_i = W_{j1} X_1 + W_{j2} X_2 + \ldots + W_{jp} X_p$$

The W_i's are known as factor score coefficients, and p is the number of variables. Detailed description of the procedure of factor analysis can be found in Norusis (1992, Chapter 2). In this research, factor analysis with the varimax rotation method was used. The purpose of the rotation is to achieve a simple, more interpretable structure.

To analyze the interrelationship among many cancer variables and their relationship with environmental factors, factor analysis was accomplished for five cancer mortality rates among both sex and race groups for the U.S. from 1953 to 1977, 1978 to 1987, and 1953 to 1987, and for Louisiana from 1953 to 1979 and from 1978 to 1987. At the county (and parish) level, 16 variables used in the factor analysis were lung, breast, colon and rectum, prostate, and stomach cancer mortality rates for both males and females and for both whites and nonwhites. Also, factor analysis was used for the analysis of a large number of environmental variables in Louisiana for the 1980s. The results of the statistical procedures are presented in sections 4.3, 5.2, and 5.4.

3.3.3 Spatial Autocorrelation Analysis

Spatial autocorrelation refers to the degree of association of a variable relative to its location. In general, if high values of a variable in one area are associated with high values of that variable in neighboring areas, the spatial pattern exhibits positive spatial autocorrelation. Conversely, when high and low values alternate at neighboring areas, the spatial autocorrelation is negative.

Spatial autocorrelation is the most fundamental element in spatial information theories, because it represents the basic relationship among features on a map. It is a central concept in geography that was discussed extensively in the 1970s and the recent developments in GIS and remote sensing have made it possible to analyze complicated spatial relationships effectively (Chou 1991; Anselin et al. 1993). A spatially autoregressive model is a generalized method and is particularly well suited to medical geographic progresses using statistical census-type data because of the inherent spatial dependencies almost always found in such data (Kennedy 1988). In previous studies on

cancer distribution patterns, spatial autocorrelation has proved to be a good measure (Glick 1977, 1979a, and 1979b; Lam 1986; Kennedy 1988; Clayton et al. 1993).

The two most useful measures of spatial autocorrelation are Moran's *I* and Geary's *C* indices. Moran's *I* index is based on the covariance of the attribute, whereas Geary's *C* index is based on the variance of the attribute. Moran's *I* is often preferred over other indices in the literature because its values follow closely our intuitive notions of positive and negative autocorrelation, and it is less affected by deviation from normality (Cliff and Ord 1973; Goodchild 1986). Moran's *I* is positive when nearby areas are similar in attributes, negative when they are dissimilar, and approximately zero when attribute values are arranged randomly and independently in space.

The formula for Moran's *I* index is as follows:

$$I = \frac{\sum_i \sum_j W_{ij} C_{ij}}{S^2 \sum_i \sum_j W_{ij}} \tag{3.6}$$

where C_{ij} represents the similarity of *i*'s and *j*'s attributes; S^2 is the sample variance; W_{ij} represents the similarity of *i*'s and *j*'s locations, and $W_{ij} = 0$ for all *i* (Goodchild 1986). In this study, W_{ij} is the sample adjacency matrix (the binary matrix) in which W_{ij} is given the value of 1 if *i* and *j* share a common boundary, and zero otherwise.

The attribute similarity measure between two areas (C_{ij}) is defined as:

$$C_{ij} = (x_i - \bar{x})(x_j - \bar{x}) \tag{3.7}$$

where x_i is the value of the attribute for area *i*, *x* bar is the mean of the attribute.

The significance of the resultant spatial autocorrelation values (from Equation 3.4) can be tested by computing *z* scores using either the normal or randomization assumption (Cliff and Ord 1973). The normal assumption states that the sample values are from a normally distributed population, whereas the randomization assumption states that the samples represent a random arrangement of attribute values.

In particular, when covariance of Moran's *I* is used, the diagram to show the relationship between spatial autocorrelation and spatial lag is called a correlogram. But when variance of Geary's *C* is used, the diagram is called a variogram. The resultant *z* scores, instead of *I* values, are plotted against spatial lags. Based on a two-tailed test with a significant level, a *z* value outside the range is considered significantly spatially autocorrelated in either a positive or negative direction. The study variable must be measured adequately and geographic units must be defined appropriately because spatial autocorrelation measures are highly scale-dependent; that is, the index changes if different sizes of polygons are used.

Using the spatial autocorrelation function in ArcInfo, this study examines the spatial-temporal patterns of clustering or spread of cancer mortality rates and analyzes the difference among sex and race-specific cancer mortality rates. Spatial autocorrelations for the five selected cancer mortality rates (lung, colon and rectum, breast, prostate, and stomach) for Louisiana from 1953 to 1987 were examined. The techniques of correlogram, autocorrelation analysis, and map presentation were utilized to investigate the spatial and temporal patterns of cancer mortality rates in Louisiana for the periods 1953-1977 and 1978-1987. Moran's *I* was used for this study because its values follow closely our intuitive notions of positive and negative autocorrelation, and

it is less affected by deviation from the normal distribution. Therefore, Moran's I was used in preference to Geary's C. The significance of resultant spatial autocorrelation values can be tested by computing the z scores using either the randomization or normal assumption. The cancer mortality data for this study were normally distributed. Therefore, the normalization assumption was more appropriate and used in the study. Using a two-tailed test with a significance level of $\alpha = 0.05$ (0.01), a z value outside the range of ± 1.96 (2.576) was considered significantly spatially autocorrelated in either a positive or negative direction.

Spatial correlograms were calculated by a computer program developed by Lam et al. (1996) (Appendix G). The computer program retrieves topological information from the ArcInfo GIS software and constructs the first-order adjacency matrix W_{ij} by coding 1 in W_{ij} if polygons i and j share a common boundary, and 0 otherwise. The higher-order adjacency matrices were derived by powering the first-order matrix and eliminating circular routes using the algorithm by Haggett et al. (1977). A description of the algorithm and procedures are described in Fan et al. (1993).

Correlogram analysis usually proceeds in four steps:

1. Extract the Arc Attribute Table (AAT) from ArcInfo to select left and right polygons that share a common boundary.
2. Covert AAT file to adjacency matrix by the C program.
3. Run the program to determine autocorrelation at each lag.
4. Draw and analyze the correlogram using Microsoft Excel.
5. The results of the statistical procedures described above are presented in section 5.3.

3.3.4 Multiple Regression Analysis

Multiple regression analysis was used to examine the relationships between several environmental variables and cancer mortality rates. Multiple regression analysis is a generalization of the simple linear regression model. Multiple linear regression extends bivariate regression by incorporating multiple independent variables. The model can be expressed as:

$$Y = a + b_1X_1 + b_2X_2 + \ldots + b_iX_i + e \tag{3.8}$$

where Y is the dependent variable, X_i are the independent variable, a is the intercept, b_i is the slope coefficients known as partial regression coefficients, and e is error term (Norusis 1992). Each partial regression coefficient represents the relationship of each corresponding independent variable with the dependent variable, while holding all other independent variables in the regression model constant. The slope coefficient can be positive, negative or zero. When the slope is positive (or negative), its contribution to the dependent variable may be positive (or negative). When it is not significantly different from zero based on an evaluation by t-test, the independent variable is said to have zero relationship with the dependent variable. The error term is the difference between the observed and expected values of Y. This model assumes that there is a normal distribution of the dependent variable for every combination of the values of the independent variables in the model.

Multiple regression is used either as a descriptive tool to summarize and decompose the linear dependence of one variable on the others or as an inferential tool by which the relationships in a population are evaluated from the examination of sample data (Goodall 1987). In particular, this model is used to seek a higher coefficient of

determination (R^2) that represents a better explanation or prediction power, and a measure of the goodness of fit of a particular model. The coefficient of determination is defined as the portion of the variance of the dependent variable explained by the independent variable(s). The larger the absolute value of the correlation coefficient, the stronger the linear association. There are several methods of calculating the regression equation but the most common is the least-squares method, which minimizes the sum of the squares of the deviations of the observed points from the fitted line (errors). F statistic (the ratio of the mean square regression to the mean square residual) is used to test regression model fit. If the probability associated with the F statistic is small, the hypothesis that $R^2 = 0$ is rejected.

For computing regression equations, there are several methods such as entry, remove, forward selection, backward elimination, and stepwise regression. None of these variable selection procedures is best in any absolute sense because they merely identify subsets of variables that for the sample, are good predictors of the dependent variable. For this study, a series of stepwise multiple regressions (using SPSS) were run on each set of variables to select which variables had the most explanatory power. This stepwise method could select good predictors, was more frequently used, and did not require as much computation.

Stepwise Selection

Stepwise selection of independent variables is probably the most commonly used method and is a combination of backward and forward procedures. It is a method that adds and removes individual variables, according to the criteria in the options dialog, until a model is reached such that no additional variables are eligible for entry or

removal. That is, the first variable is selected in the same manner as in forward selection. This procedure begins by building a model with the single variable that explains the most variation in the dependent variable. Additional variables are added in order of their contribution to the overall R^2, provided they meet a predetermined level of significance. After a variable is added to the equation, the variables that have already entered are tested against the predetermined significance level to determine whether they should be removed from the model (as in backward elimination). Variable selection terminates when no more variables meet entry and removal criteria. Variables can be entered or removed from the model depending on either the significance of the F value, or the F value itself. Entry and removal values must be greater than 0 and less or equal to 1, and the entry value must be less than the removal value. This study used 0.05, the probability associated with the F statistic to enter and 0.1, the probability associated with the F value for removal.

The weakness of the forward selection lies in the fact that it only tests variables one at a time for entry into the model. Sometimes, several independent factors together will explain a significant amount of the variation in the dependent variable, but, if they are each tested separately, none will meet the criterion of significance to enter the model. To explore that possibility, a backward elimination was run on the same set of variables. The backward technique initially determines a model based on the entire set of variables. Individual variables are then considered in order of their contribution to the overall R^2, with the smallest contributors considered first. Variables that are not significant at the smallest level allowed are removed from the model (Nickerson 1978).

Forward selection has some advantage over the backward procedure because it tests both for entry into the model and for continued significance once the other variables have entered. In backward elimination, once a variable is removed, it is not considered for re-inclusion in the model.

For this study, stepwise regression analysis was applied to examine the relationships between 24 environmental variables and cancer mortality rates for the 1980s. The results are presented in section 5.5.

3.3.5 Spatial Scan Statistic Analysis

Disease cluster detection is important to the study of distribution, possible causes, and control of disease, as it may ultimately lead to deviation of etiologic factors. A scan statistic is commonly used to test if a one-dimensional point process is purely random, or if any clusters can be detected. The general statistical theory for the scan statistic is described in detail by Kulldorff (1997). Turnbull et al. (1991) developed a generalization of a test for the spatial scan statistic. Application of the scan statistic can be found in Kulldorff et al. (1995), Hjalmars et al. (1996), and Kulldorff et al. (1997).

A scan statistic uses two different probabilistic models, the Bernoulli and Poisson distributions. With the Bernoulli model, there are cases and non-cases as a 0/1 variable. These may represent people with or without a disease, or people with different types of diseases. On the other hand, for the Poisson model, the number of cases in each census area is assumed to be Poisson distributed. With either model, the scan statistic adjusts for the uneven population density present in almost all populations, and the analysis is conditioned on the total number of cases observed.

The spatial scan statistic imposes a circular window on the map. The window is in turn centered around each of several possible centroids positioned throughout the study area. For each centroid, the radius of the window varies continuously in size from zero to some upper limit so that the window never includes more than 50 percent of the total population. In this way, the window is flexible both in location and size. In total, the method creates a very large number of distinct geographical circles, with different sets of neighboring census areas within them, and each being a possible candidate for a cluster (Kulldorff et al. 1997). The temporal scan statistic has a window that moves in one dimension (time), defined by a cylindrical window with height corresponding to time. The space-time scan statistic is defined by a cylindrical window with a circular geographic base and with height corresponding to time. The base is defined exactly as for the purely spatial scan statistic, while the height reflects the time period of potential clusters. The cylindrical window is then moved in space and time, so that for each possible geographical location and size, it also visits each possible time period (SaTScan 2.1 1998).

For each location and size of the scanning window, this method tests the null hypothesis against the alternative hypothesis that there is an elevated rate within the window as compared to outside. This study assumed the number of deaths in each parish or census tract to be Poisson distributed. Under the Poisson assumption, the likelihood function for a specific window is then proportional to

$$\left(\frac{n}{u}\right)^{n}\left(\frac{N-n}{N-u}\right)^{N-n} I(n>u) \tag{3.9}$$

where N is the total number of cases (deaths) over the whole area, n is the number of cases within the window, and u is the covariate adjusted expected number of cases within the window under the null-hypothesis. I is an indicator function that is equal to 1 when the window has more cases than expected under the null-hypothesis and 0 otherwise (SaTScan 2.1 1998).

The likelihood function is maximized over all windows, identifying the window that constitutes the most likely cluster. The likelihood ratio for this window is noted and constitutes the maximum likelihood ratio test statistic. Its distribution under the null-hypothesis and its corresponding p-value is obtained by repeating the same analytic exercise on a large number random of replications of the data set generated under the null hypothesis, in a Monte Carlo simulation. The method also identifies secondary clusters according to the likelihood ratio. The secondary clusters may or may not overlap the most likely cluster (Kulldorff et al. 1997).

In addition to scan statistic, there are many different statistical methods for disease clustering, such as descriptive cluster detection, focused, global clustering, and space-time interaction test.

Openshaw et al. (1987) have developed a Geographical Analysis Machine (GAM) that uses overlapping circles of different sizes in the same way as the spatial scan statistic, except that the circle size does not vary continuously. Also, the cluster detection method proposed by Rushton and Lolonis (1996) is similar to that of Openshaw et al. Both methods are very useful for descriptive purposes, but should not be used for hypothesis testing.

Focused tests are used to examine the presence of an elevated risk of the disease around the specific source when there is a prespecified point source. These studies include Bithell's test (1995), the Lawson-Waller score test (1993), Stone's test (1988), and isotonic binary regression.

Global clustering is one of the most studied tests for spatial clustering of health events. This includes the methods proposed by Alt and Vach (1991), Besag and Newell (1991), Cuzick and Edwards (1990), Diggle and Chetwynd (1991), Grimson (1991), Moran (1950), Ranta et al. (1996), Tango (1995), Walter (1994), and Whittemore et al. (1987). These methods test for clustering throughout the study region without the ability to pinpoint the location of specific clusters. As such, these tests and the spatial scan statistic complement each other very well, being useful for different purposes.

Space-time interaction methods have been proposed by Knox (1964), Mantel (1967), Diggle et al. (1995), Jacquez (1996), Baker (1996), and Kulldorff and Hjalmars (1999). These methods are designed to evaluate whether cases that are close in space are also close in time and vice-versa, adjusting for any purely spatial or purely temporal clustering. They are very useful when testing to see if there is clustering throughout the study area and time period, and preferred methods when, for example, trying to determine whether a disease is infectious. However, unlike the space-time scan statistic, they are unable to detect the location and size of clusters and to test the significance of those clusters.

Compared with other statistical methods for spatial epidemiology, the spatial scan statistic has the following features that make it particularly suitable as a screening tool for evaluating reported disease clusters:

1. By searching for clusters without specifying their size or location, the method ameliorates the problem of preselection bias.

2. It adjusts both for the inhomogeneous population density and for any number of confounding variables.

3. The likelihood ratio-based test statistic takes multiple testing into account and delivers a single *p*-value for the test of the null hypothesis.

4. If the null hypothesis is rejected, we can specify the approximate location of the cluster that caused the rejection (Kulldorff et al. 1997).

Therefore, this dissertation employed the SaTScan software (2.1), which analyzes the point data using the spatial, temporal, or space-time scan statistic, to detect and compare the geographical clustering of lung cancer in Louisiana by parish and by tract from 1988 to 1993. The data for this analysis are in three different files: geographical coordinates (latitude and longitude), case counts, and population. Also, a scan statistic by adjusting covariates such as age, sex, and race was run with the Poisson model (eg., adjusting covariate sex means selecting only male or female cases). The SaTScan software has been integrated into Arcview, and special output files describing the various clusters can be displayed and analyzed through Arcview. These files can also be accessed using any text editor or spreadsheet program. The results of the statistical procedures described above are presented in Chapter 6.

3.4 Limitations of the Study

This research should consider at least five methodological problems.

First, availability and quality of data should be considered. For example, studies of cancer incidence may be more suitable than those of cancer mortality statistics in

order to identify the possible causal factors of diseases. But unlike mortality data, incidence data for all parishes have been published for limited short periods. Although TRI data is the most comprehensive data currently available on potential environmental exposure to toxic chemical releases, it includes only manufacturing facilities and the data-collection system relies on self-reporting and permits estimation. Therefore, the toxic chemical emissions do not necessarily reflect actual discharges of total emissions.

Second, many chronic and infectious diseases exhibit time lags between exposure to causal agent and the initial detection of disease symptoms, making detection of causal agent more difficult.

Third, it is difficult to address the causal relationships between mortality and environment, because identification of such relationships is hampered by migration in highly mobile societies such as this study remains.

Fourth, there is a set of issues and problems relating to both the temporal and spatial scale of a medical-geographical study. The issues of ecological fallacy and scale problems which are central to spatial aggregation studies such as this study remain (Openshaw 1977).

Fifth, genetic factors as etiological roles in cancer are still not clear. More studies on the interplay between genetic factors and environmental factors are needed.

3.5 Summary

This chapter contains a description of the research designs, such as study area, data source, statistical methods, and limitations of the study, to examine the relationship between cancer mortality rates and environmental factors. Study areas consisted of the U.S and Louisiana, at the state, county (parish), and census tract levels. Data on cancer

(from the1950s through the 1990s) as well as environmental variables (in the 1980s) in Louisiana, and cancer mortality rates in the U.S. (from 1953 to 1987) were collected. Table 3.2 summarizes the data by geographic scale and time period. The data were analyzed according to the three hypotheses of the study. The methodologies of this research are summarized in a flowchart (Figure 3.4).

In the next chapter, the results of the analyses are reviewed, and the conclusions of hypotheses are presented.

Table 3.2 Summary of Data (Bark, 1999)

DATA	AREA		PERIOD
Cancer: Cancer Mortality	U.S. (state & county)	Lung, breast, colon & rectum, prostate, and stomach	1953-1987
	LA (parish)	Lung, breast, colon & rectum, prostate, and stomach	1953-1987, 1980-1989, 1988-1993
Cancer Incidence	LA (parish)	Lung, breast, colon & rectum, prostate, and stomach	1983-1987, 1988-1993
Death Number	LA (parish & tract)	Lung, breast, colon & rectum, prostate, and stomach	1988-1993
Environmental Variables: (LA)	(parish) (tract)	Population Population	1988-1993 1980-1993
	(parish)	Population (White & Nonwhites)	1980, 1990
		Land Area	1980,1990
		Population Density	1980,1990
		Per Capita Income	1980-1989
		Education Status (High school graduates)	1980,1990
		Persons below poverty level	1979,1989
		Occupation (Agriculture, construction, manufacturing-all products, chemical manufacturing, mining, transportation, health services, and educational services	1980-1989
	(parish) (tract)	TRI releases (Air, water, ground injection, and on-site land)	1987-1989
	(parish)	Waste sites (Solid, inactive & abandoned, and hazardous waste sites)	1980-1989
	(parish)	Pesticide (Use in agriculture)	1982,1987
	(parish) (tract)	Wetlands	1978,1987
	(parish)	Urban population	1980,1990
	(parish)	Air quality (Ozone)	1980-1990
	(parish)	Water quality (Surface water or Mississippi River)	1980-1989
Geographic Location:	(parish) (tract)	Latitude, Longitude Latitude, Longitude	1988-1993 1988-1993

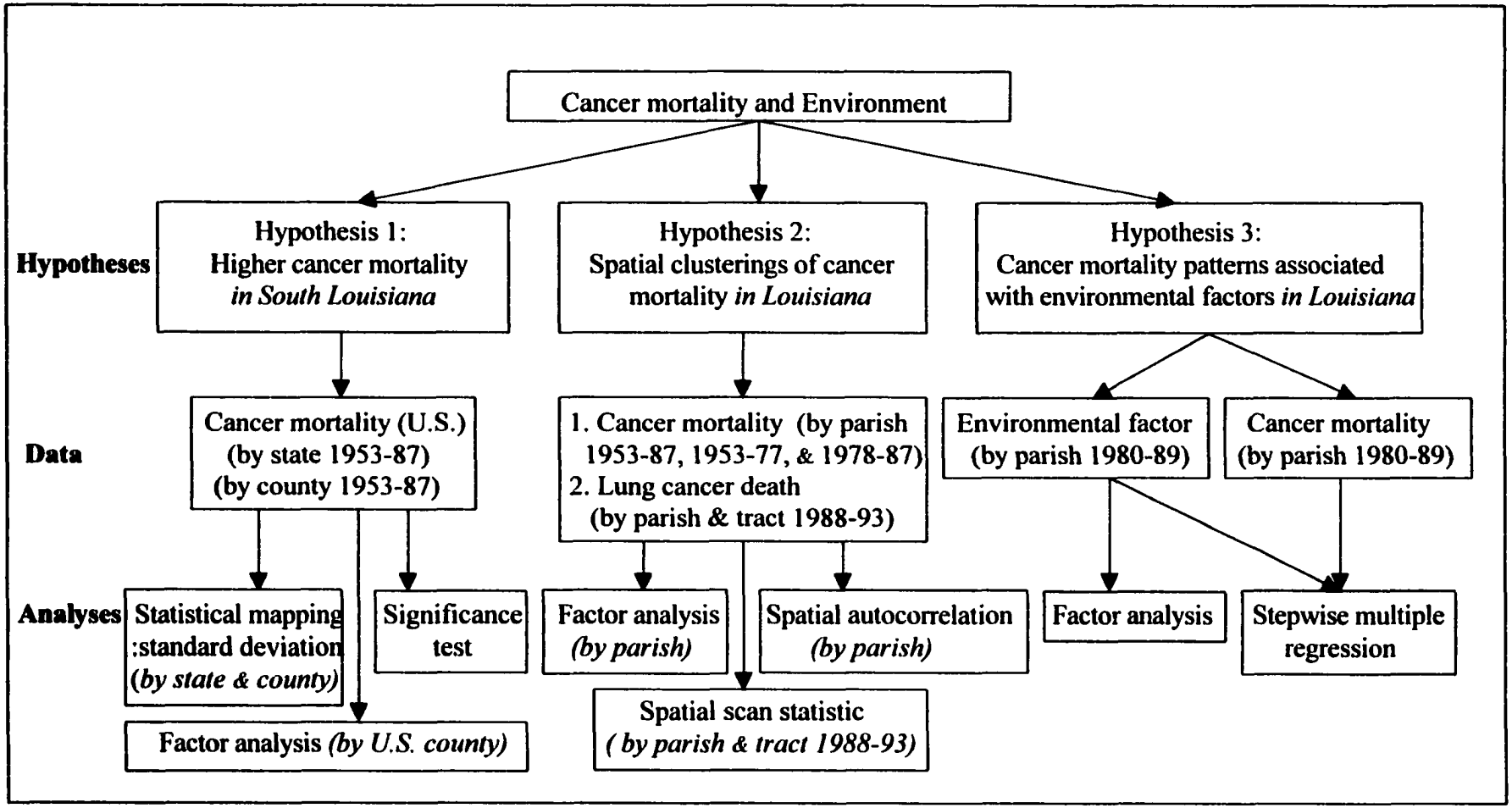

Figure 3.4 Flowchart of Methodology

CHAPTER 4

NATIONAL PATTERNS

The purpose of this chapter is to provide a base line of cancer patterns in the U.S. during the 1953 through 1987 study period. It is used also to examine the first hypothesis--cancer mortality rates are higher in South Louisiana than in the nation and the state. The mapping by standard deviation classification and the factor analysis method outlined in Chapter 3 were used to generate the results presented below. This chapter has three components. The first and second are discussions of the geographic patterns of cancer mortality rates by state and county in the contiguous U.S., including the District of Columbia, for the period 1953 to 1987. The third component discusses the results of factor analysis for the five major cancer mortality rates in U.S. counties from 1953 to 1987.

4.1 Geographic Patterns of Cancer Mortality Rates in the U.S. by State

In examining the cancer mortality patterns by state from 1953 to 1987, the age-adjusted cancer mortality rates (Appendix H) were mapped using the standard deviation method (Figures 4.1- 4.5). There were wide variations in total or specific cancer mortality rates among different states. High cancer mortality rates were markedly shown in the northeastern U.S. The average annual age-adjusted cancer rates of all sites combined for the entire U.S. during the study period were 164.2 deaths per 100,000 population, and rates of males, females, whites, and nonwhites were 202.8, 136.1, 161.7, and 185.6 deaths per 100,000 population, respectively.

Figure 4.1 shows the total death rates from cancer of all sites combined over the thirty-five year period. The northeastern U.S. (the District of Columbia, Maryland, New

Figure 4.1 Cancer Mortality Rates for All Sites: 1953-1987
(Source: Calculated by Author from National Technical Information Service, 1992)

Jersey, Delaware, Rhode Island, New York, New Hampshire, Pennsylvania, Massachusetts, Connecticut, Ohio, Maine, Illinois, and Michigan), Louisiana, and Nevada had high cancer mortality rates. Mountain regions (such as Utah, Idaho, Colorado, Wyoming, and New Mexico) had low cancer mortality rates. These low mortality rates extended to West North Central regions (such as North Dakota, Kansas, and South Dakota). The District of Columbia had the highest rate among the 49 states, 205.7 deaths per 100,000 population for the thirty-five year period, whereas Utah experienced the lowest, 122.6 deaths per 100,000.

Figures 4.2 and 4.3 display the cancer mortality rates for males and females. The northeastern U.S. and Louisiana had high rates whereas Mountain regions except for Nevada had low rates among males (Figure 4.2). The District of Columbia had the highest rates for males, 272.9 deaths per 100,000 population, and Utah had the lowest, 147.3 deaths per 100,000. Figure 4.3 shows that the states with high rates in females were all located in the northeastern U.S and Nevada, whereas the states with low rates for females were distributed in Mountain and South regions. Similar to the case of male deaths, the District of Columbia had the highest female death rate, 162.1 deaths per 100,000 population, and Utah had the lowest, 104.3 deaths per 100,000. Comparison of Figures 4.2 and 4.3 reveals that males have higher cancer death rates than females in every state. Male to female (M:F) ratios are approximately 2:1.

Figures 4.4 and 4.5 display the cancer mortality rates among whites and nonwhites. Among whites, all the states in the northeastern U.S. fell well above the mean, while other states fell well or slightly below the mean except for Nevada and Louisiana. Nonwhite cancer mortality rates were predominantly high in the northeastern

Figure 4.2 Cancer Mortality Rates for All Sites (Males): 1953-1987
(Source: Calculated by Author from National Technical Information Service, 1992)

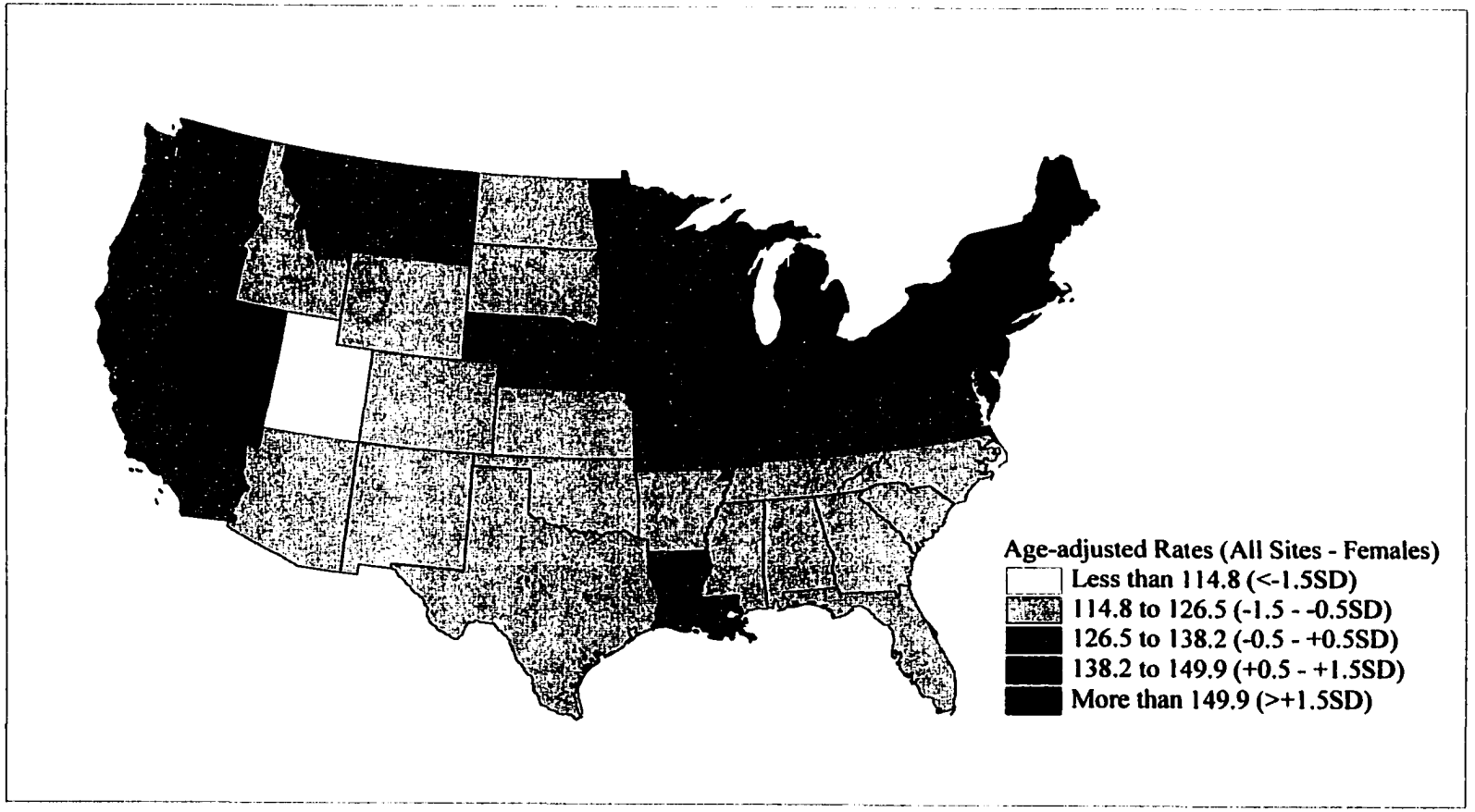

Figure 4.3 Cancer Mortality Rates for All Sites (Females): 1953-1987
(Source: Calculated by Author from National Technical Information Service, 1992)

Figure 4.4 Cancer Mortality Rates for All Sites (Whites): 1953-1987
(Source: Calculated by Author from National Technical Information Service, 1992)

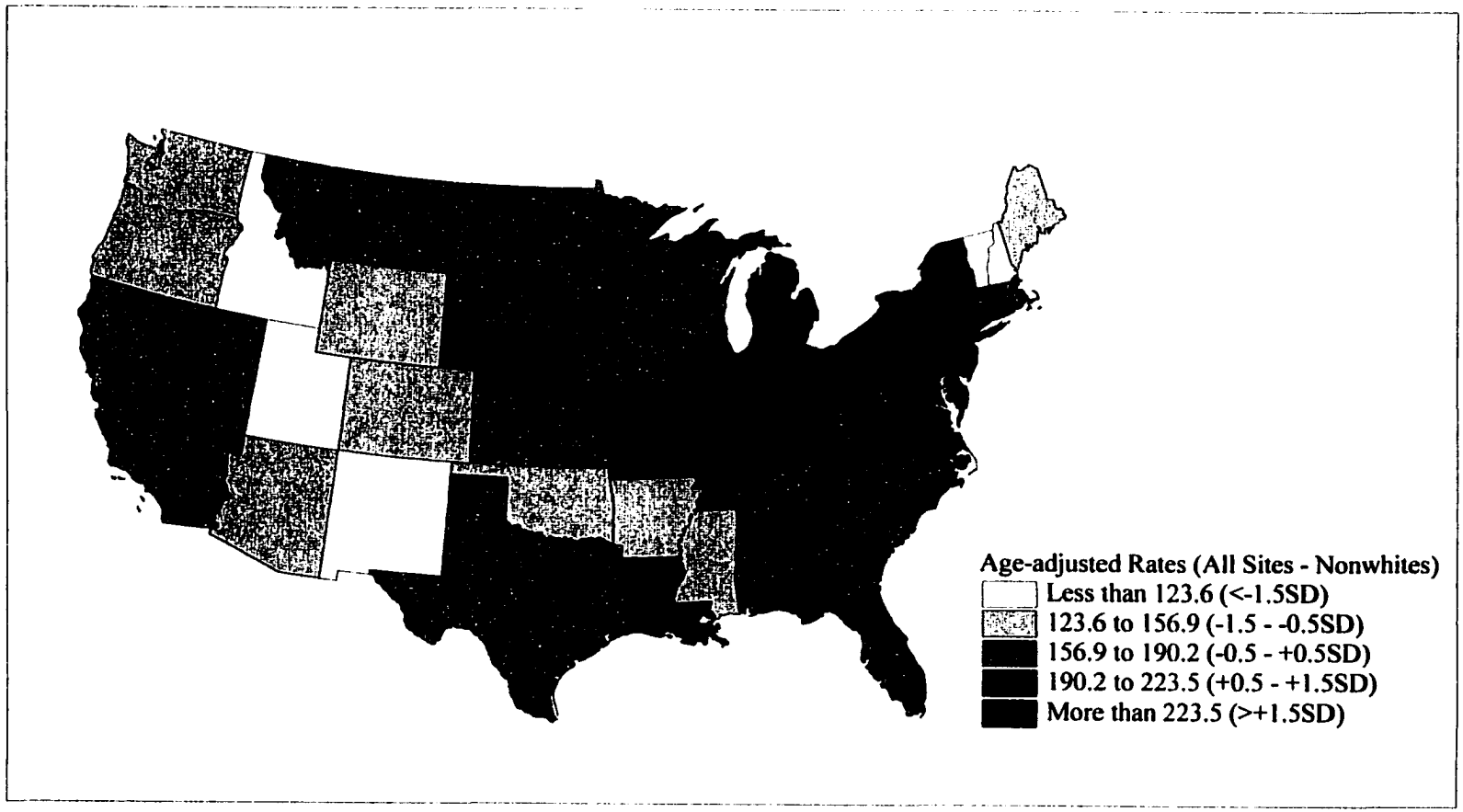

Figure 4.5 Cancer Mortality Rates for All Sites (Nonwhites): 1953-1987
(Source: Calculated by Author from National Technical Information Service, 1992)

U.S. (except for New England regions), Nebraska, and Louisiana. The highest rates were observed for whites in New Jersey (181.3) and for nonwhites in the District of Columbia (236.2). The lowest rates were for whites in Utah (122.2) and for nonwhites in New Hampshire (89.7). In most states, nonwhites had higher cancer rates than whites. However, New Hampshire, Vermont, and Maine in New England regions, New Mexico, Idaho, and Arizona had higher cancer mortality rates among whites than nonwhites.

In this section, cancer mortality rates in the U.S. states were briefly compared to determine what regional variations existed in the spatial distributions of different kinds of cancer. Several observations could be made. Males had higher cancer rates than females of the same race. Cancer occurred more frequently among nonwhites. The spatial patterns between males and females for the same cancers were generally very similar, whereas those between whites and nonwhites were different.

The northeastern U.S. during 1953-1987 showed high cancer mortality rates whereas the mountain states had relatively low rates. The most striking feature of this cluster was that the highest incidence states were in the northeastern part of the country. The District of Columbia was found to have the highest mortality rates from 1953 to 1987. This state has a higher percentage of blacks (about two-thirds) than other states and is predominantly an urban area whereas other states are composed of a combination of urban, suburban, and rural areas. Another somewhat higher than average regional cluster was situated in the southwestern U.S. (such as Nevada and California). In addition, mortality statistics and maps at the state level have consistently shown high rates in Louisiana, which is discussed in Chapter 5 in detail. The age-adjusted cancer rates of all sites combined for Louisiana from 1953 to 1987 were 173.0 deaths per

100,000 population, and rates of males, females, whites, and nonwhites were 225.1, 134.2, 166.6, and 191.1 deaths per 100,000 population, respectively.

4.2 Geographic Patterns of Cancer Mortality Rates in the U.S. by County

For the cancer mortality patterns by county from 1953 to 1987, the age-adjusted cancer mortality rates were mapped using the same standard deviation method for the state level. The county level data showed wider variations in cancer mortality rates among different counties than the state level. The areas that generated high cancer mortality rates at the state level also showed relatively high rates at the county level. The age-adjusted cancer mortality rates of all sites combined for the entire U.S. from 1953 to 1987 were 152.7 deaths per 100,000 population, and rates of males, females, whites, and nonwhites were 186.5, 126.5, 153.1, and 175.3 deaths per 100,000 population, respectively. For cancer in all sites combined, the high mortality rate areas were generally concentrated in the northeastern seaboard megalopolitan area (Figure 4.6). Parts of the Midwest also displayed high rates. Coastal Louisiana, Florida in the South, and California and Nevada in the western half of the country were the areas where high rates consistently developed. Most of the West, including substantial portions of the Great Plains, displayed only a few dispersed areas of high rates, and large sections generated modest to low rates.

During the 35 years, the cancer sites with high mortality rates were cancers of lung, prostate, and colorectum among males and of breast, lung, and colorectum among females. The geographic distributions for the most commonly occurring cancers from 1953 to 1987 are described in Figures 4.7 - 4.11. The patterns of cancer mortality rates

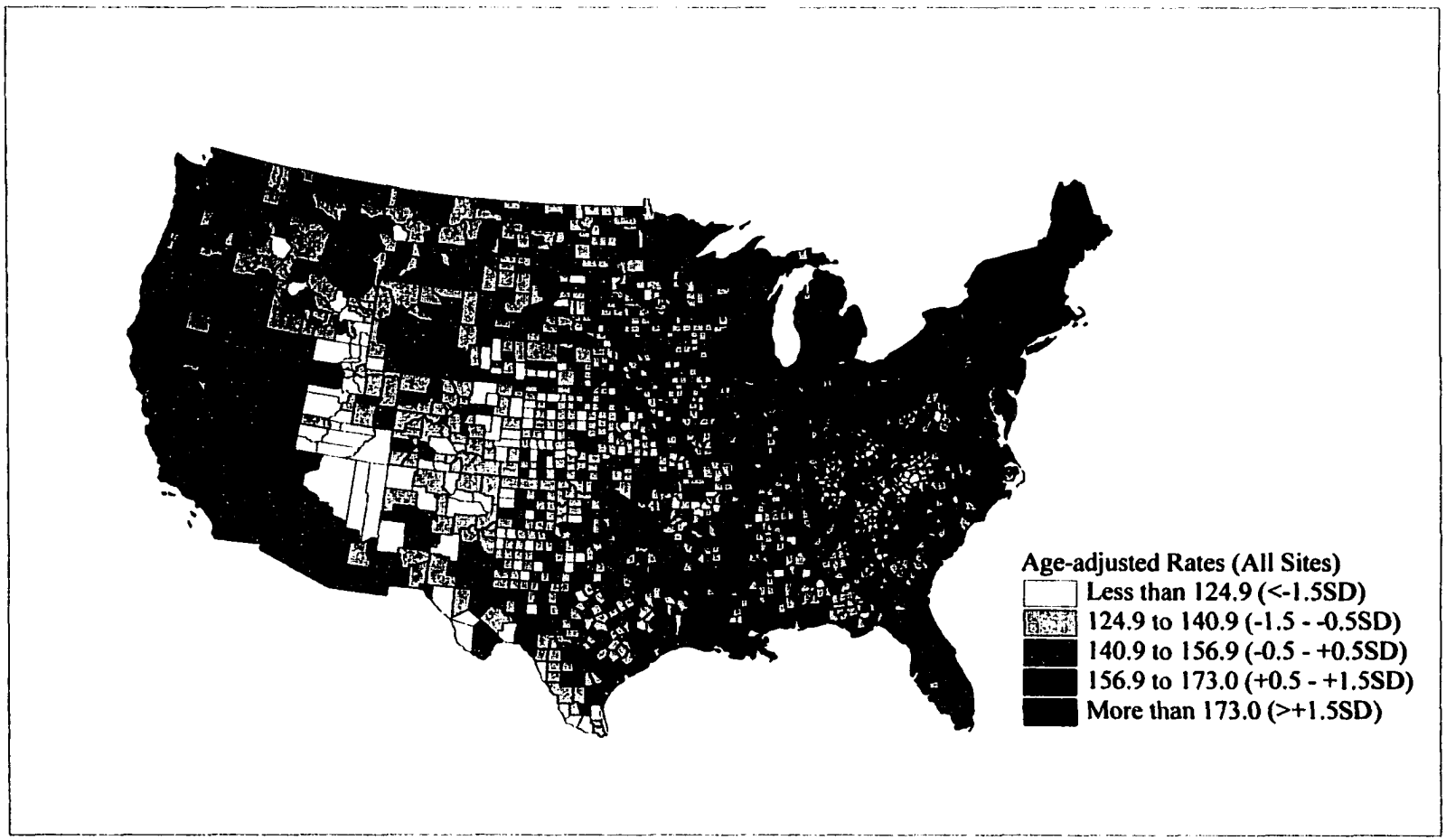

Figure 4.6 Cancer Mortality Rates for All Sites (by County): 1953-1987
(Source: Calculated by Author from National Technical Information Service, 1992)

were summarized in the following (Note that for convenience of this study, only the total cancer death rates for 1953-1987 were mapped).

4.2.1 Breast Cancer

Breast cancer accounts for more deaths among females than any other type. From the data, it seems that the geographic patterns have remained generally stable over time and there has been a consistent North-South gradient (Figure 4.7). In general, high mortality rates for breast cancer were concentrated in the northeastern quadrant of the country. For instance, a clustering of the highest rate areas was developed in the urban Northeast, although scattered high rate areas appeared in some rural counties. Another concentration was shown in counties of upper Midwest. Exceptionally high rates were found in Camas, Idaho (74.11); Thomas, Nebraska (43.03); and Petroleum (42.37) and Wheatland (42.21), Montana. Most of the West coast showed average to higher than average rates for breast cancer mortality. The low rates for breast cancer mortality rates were shown in the southern half of the country and some counties of Texas, Utah, Arizona, and New Mexico states. The lowest rates were in Borden, King, and Loving, (0.00) Texas; Sterling, Texas (4.33); and Union, Florida (4.98).

The geographic pattern of breast cancer mortality rates for white females were very similar to those for all females. As for nonwhites, the counties with the high mortality rates were scattered in the lower Northeast and centers of the Midwest. A broad area that consisted of most of the South has displayed rising mortality rates and the North-South differences have diminished, when compared with the geographic distributions of breast cancer mortality rates for overall or white females.

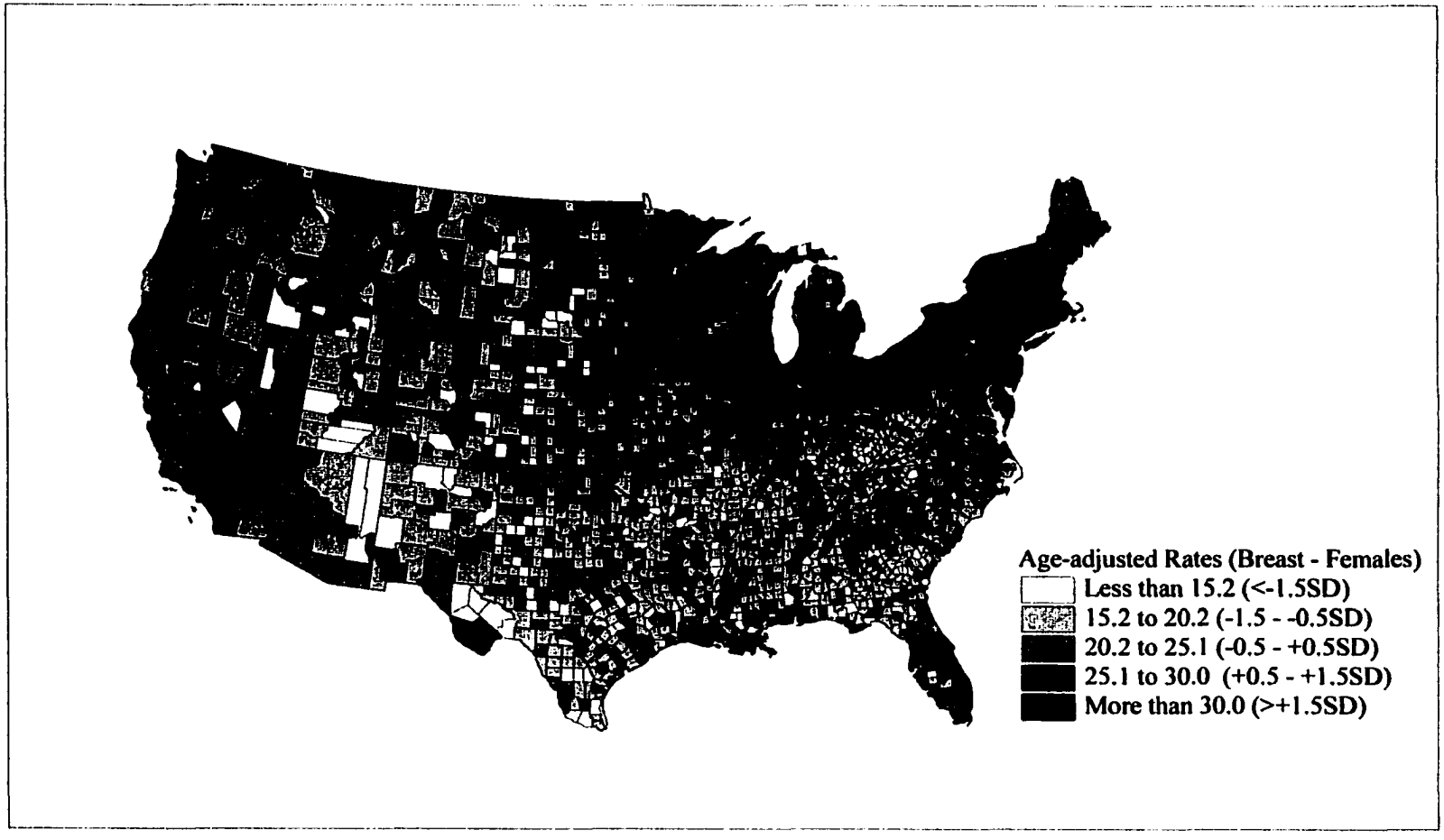

Figure 4.7 Breast Cancer Mortality Rates (All Races, Females): 1953-1987
(Source: Calculated by Author from National Technical Information Service, 1992)

4.2.2 Colorectal Cancer

Colorectal cancer means the cancers of colon and rectum and the geographic patterns for colon cancer resemble those for rectal cancer (Pickle et al. 1990). There have been more than twice as many deaths of colon cancer than of rectal cancer (Correa et al. 1983a). As in the case of breast cancer, areas with the highest mortality rates were in the northeastern quadrant of the country, particularly in counties of New Jersey, Massachusetts, New York, southern Maine, and Pennsylvania (Figure 4.8). In addition to this cluster, Loving (44.21) and King (42.09), Texas; San Juan (36.89), Colorado; and Stanley (35.18), South Dakota showed exceptionally high mortality rates. Low death rates were markedly concentrated in the southern half of the country, such as counties of the South and South Mountain regions. Counties with the lowest rates were in Kenedy (0.00) and Irion (5.09), Texas; San Juan (5.65) and Wayne (6.06), Utah, respectively. Even though it was not shown in the maps, there was a consistent North-South gradient of spatial distribution in the early years, its difference has diminished with time, as many areas in the South have increased to have higher mortality rates than in the North.

Similar to the overall rates, the spatial patterns for males, females, and whites showed high mortality rates in the northern parts of the U.S. However, colorectal cancer mortality rates for nonwhites generated a concentration of high rates in Pennsylvania and Ohio. Also, some counties in West Virginia, Kentucky, Virginia, Tennessee, and western North Carolina showed high rates. Most counties of the West and South developed moderate rates and most of the West North Central regions showed lower rates of colorectal cancer mortality.

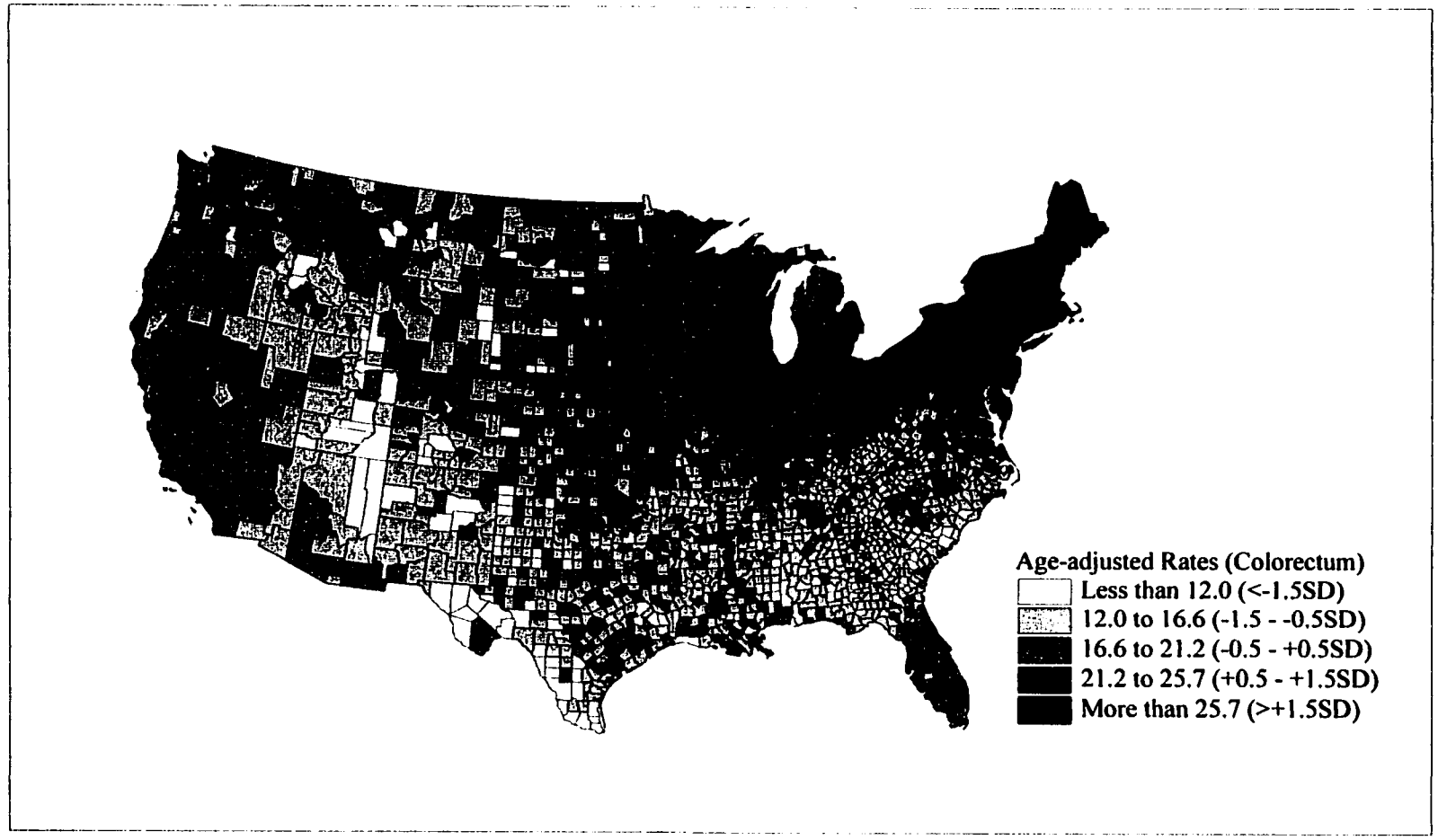

Figure 4.8 Colorectal Cancer Mortality Rates (All Races, Both Sexes): 1953-1987
(Source: Calculated by Author from National Technical Information Service, 1992)

The spatial distributions of colorectal cancer mortality rates were similar to those of breast cancer. The distributions of high rates for colorectal cancer were a little more concentrated than those of the breast cancer while areas of low rates were a little more fragmented than those of the breast cancer.

4.2.3 Lung Cancer

As was mentioned before, lung cancer is the leading cause of cancer death. The most region of high lung cancer mortality rates prevailed in the southeastern U.S. (Figure 4.9). Core high-rate counties for the region during the 1953-1987 were Okeechobee (56.95), Union (56.90), Glades (54.37), and Franklin (53.67) in Florida, Reagan (54.77) in Texas, Charlton (54.44) in Georgia, and St. Bernard (53.43) and St. Tammany (51.23) in Louisiana, with much higher than average rates also in many adjacent counties. Secondary areas of high lung mortality rates emerged in the northeastern part of the country and the West coast. In the western part of the U.S., most counties in California, Oregon, and Washington reported somewhat higher than average figures, along with some counties in Nevada, western Arizona, and Grand in Utah. Areas of low rates were dispersed in the Midwest and Mountain region.

Lung cancer occurs more often among men than among women and more often among nonwhites than among whites. Geographical patterns of lung cancer mortality rates were the same in the 1980s as in the 1950s, but there were also some major changes (not shown in the map). Lung cancer mortality rates in the Southeast have risen to lessen the North-South differences. The spatial pattern of lung cancer mortality rates was different from the previous two distributions (breast and colorectum). Concentrations of much higher than average cancer death rates were in the southeastern

Figure 4.9 Lung Cancer Mortality Rates (All Races, Both Sexes): 1953-1987
(Source: Calculated by Author from National Technical Information Service, 1992)

U.S. While the disease is much more widespread among males, the death rates have also been on the increase among females in recent years. The average of male lung cancer mortality rates was 3 times higher than that of female rates during the period 1953-1987.

The mortality rates among males were high in certain seaboard areas of the Southeast Atlantic and Gulf coasts, especially along counties extending from the southeastern Texas and South Louisiana through Florida and South Georgia. The mortality rates has declined in the Northeast and increased in the South. The central part of the U.S. continued to show lower than average lung cancer mortality rates for males, with Utah reporting the lowest rate.

High concentrations of female lung cancer mortality rates occurred in the southeastern U.S. and the West. Most counties of Nevada, California, and Florida showed high rates for the period 1953-1987. Some northeastern counties also showed average to higher than average rates. The rate of increase among females has risen sharply throughout the country, with concentrations of high rates in Florida and along the mid-Atlantic and West coast.

The geographic patterns of lung cancer mortality rates among whites were very similar to those of overall lung, with three areas (southeastern, northeastern, and the West Coast) displaying high rates. Among nonwhites, the distributions of lung cancer mortality rates differed considerably from those among whites. The highest rates for nonwhite lung cancer mortality were widespread and did not show a strong concentration. Clusters of relatively high rates for the cancer mortality were shown in southern Louisiana, Florida, and areas scattered throughout the northern half of the U.S.

Areas of low rates were much larger and more widely dispersed throughout the Midwest and Mountain region.

4.2.4 Prostate Cancer

Prostate cancer is the second most common cancer among males in the U.S. The geographic patterns of prostate cancer mortality rates were totally different from those of other cancer sites (Figure 4.10). Areas of the highest mortality rates of prostate cancer occurred in counties of South Carolina, North Carolina, and Utah, whereas the lowest rates were found in counties of Nevada and Arizona. These clusterings were relatively small and widely dispersed in the U.S. In particular, counties of the highest mortality rates included Loving, Texas (89.6); Grant, Nebraska (75.7); and Petroleum, Montana (48.3). The counties of Daggett (Utah), King (Texas), Mcpherson (Nebraska), Clark (Idaho), and San Juan and Hinsdale (Colorado) showed zero, the lowest rates.

Areas of high mortality rates of prostate cancer were much smaller and more widely distributed in the southern two-thirds of the country. Areas of low cancer rates included most of the southern West regions, the mountain area of Kentucky, West Virginia, and Florida.

In white males, although little geographic variation for prostate cancer mortality rates has been apparent, some clusterings of high rate areas occurred in some counties of the northern Mountain and West North Central regions and the Northeast. Areas of low cancer rates were scattered throughout the southern half of the U.S. High rate areas for prostate cancer mortality among nonwhite males concentrated in the southern one-third of the country, such as most of the South Atlantic region and some counties of

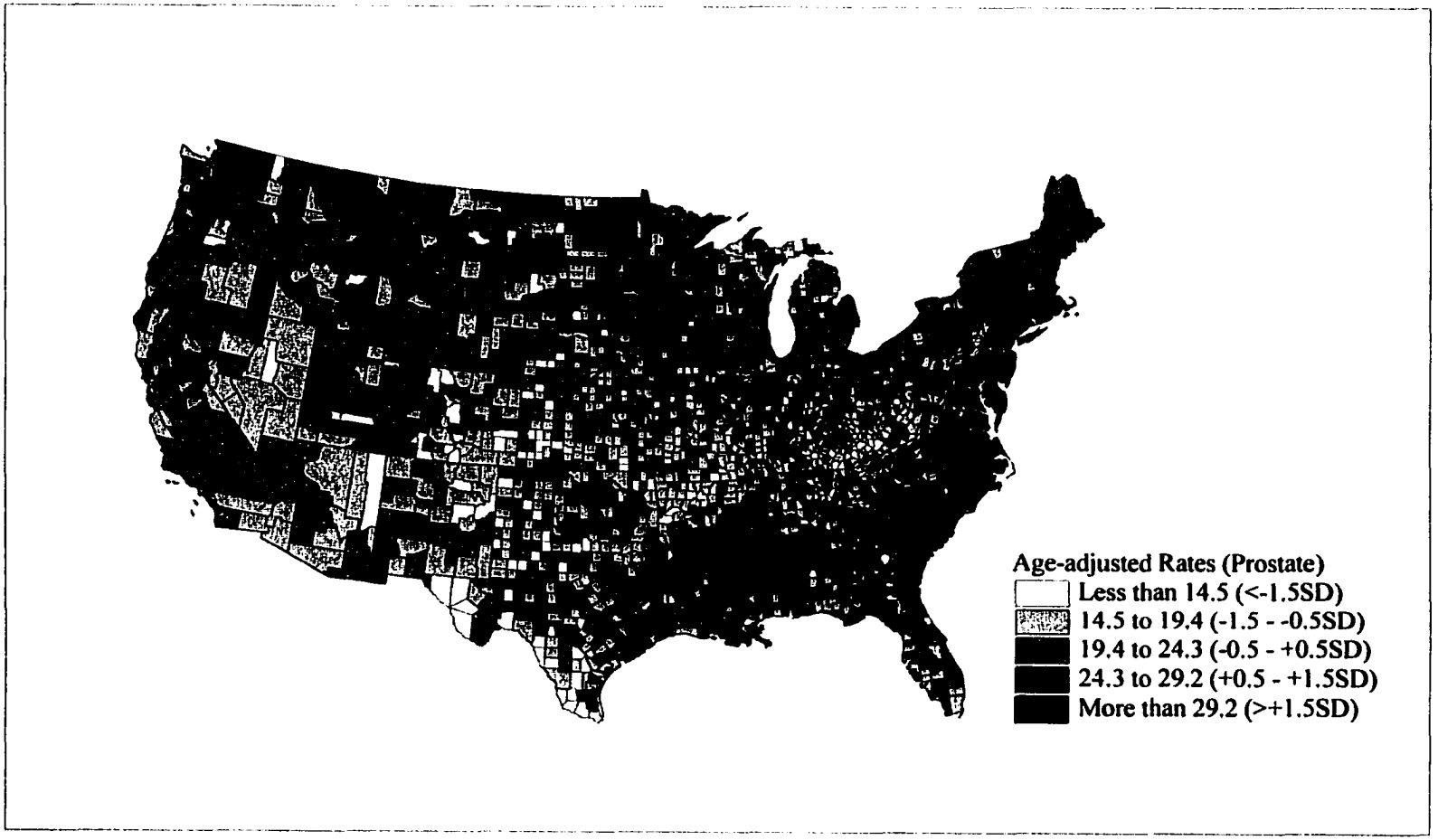

Figure 4.10 Prostate Cancer Mortality Rates (All Races, Males): 1953-1987
(Source: Calculated by Author from National Technical Information Service, 1992)

Louisiana and western Texas. Areas of low cancer mortality rates were generally distributed in the West and Midwest.

4.2.5 Stomach Cancer

Stomach cancer was the leading cause of cancer death among both sexes in the U.S. in the 1930s. But there has been a rapid decrease in stomach cancer mortality rates and the U.S death rates for stomach cancer have been among the lowest in the world.

Areas of high rates of stomach cancer mortality developed in the upper Midwest and northern New Mexico, not in the Northeast (Figure 4.11). Secondary areas of high stomach rates emerged in southern Louisiana, except Cameron parish. Counties with the highest mortality rates included Loving, Texas (31.10), Summit, Colorado (30.20), Esmeralda, Nevada (22.04), Burke (18.14) and Logan (17.49), North Dakota, and Guadalupe, New Mexico (17.30). The counties of Daggett (Utah), Hinsdale and Mineral (Colorado) showed the lowest rates, zero.

Areas of low rates were dispersed in the West and the southern one-half of the country. The leading regions were the lower Midwest region, Wyoming, eastern Texas, and peninsular Florida. Generally, there were more areas of higher than average rates in the western U.S than in the eastern counties and the northern counties than the southern counties.

Rates among males and nonwhites were almost twice as high as those among females and whites, respectively, for cancers of the stomach. Regional patterns in the mortality rates of white stomach cancer were somewhat similar to those of overall stomach cancer, except for Louisiana parishes. Stomach cancer mortality rates for whites among both sexes were very low in the Southeast. However, New Mexico and

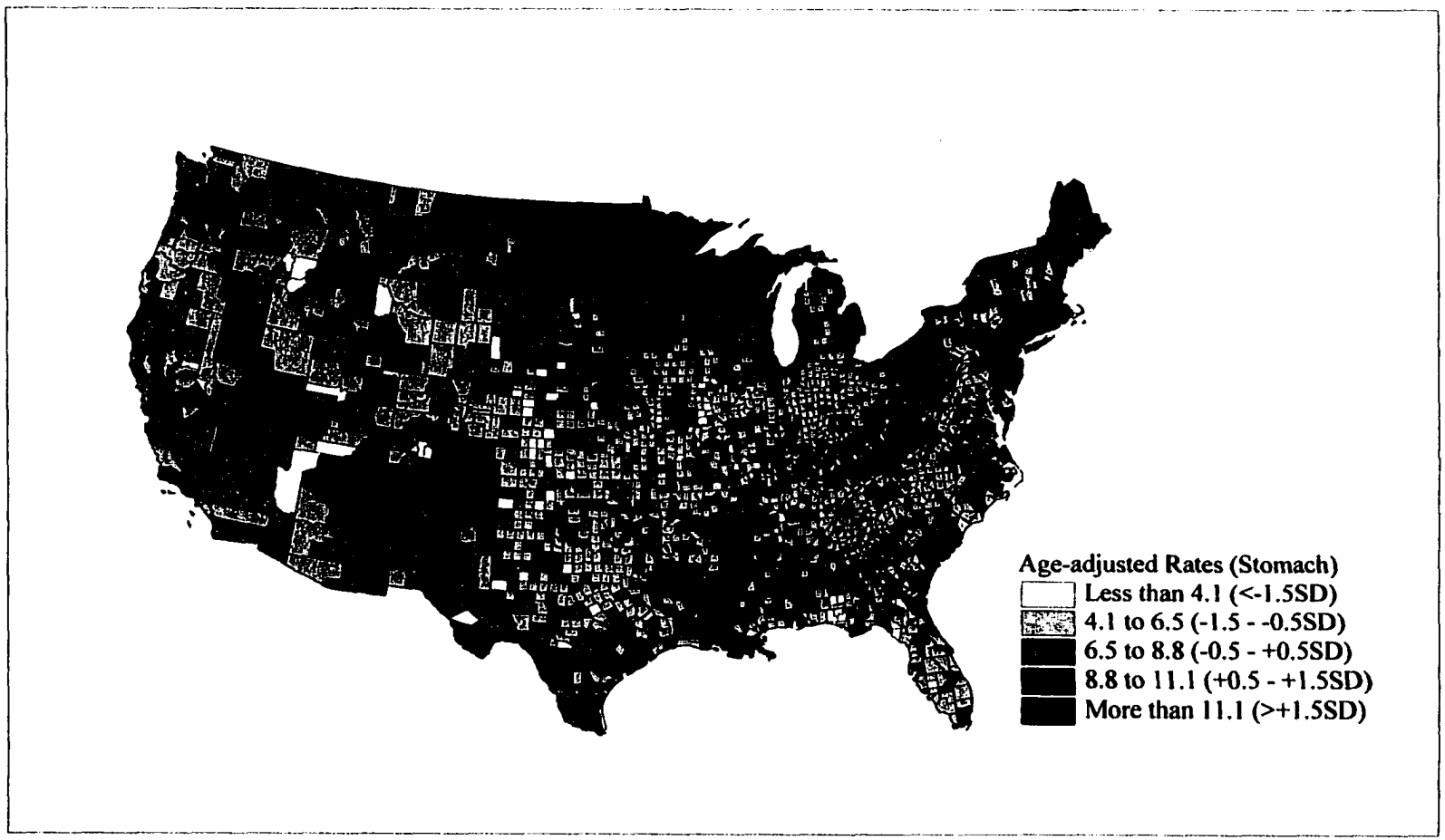

Figure 4.11 Stomach Cancer Mortality Rates (All Races, Both Sexes): 1953-1987
(Source: Calculated by Author from National Technical Information Service, 1992)

North Dakota stood out with exceptionally high mortality rates, along with parts of the Northeast.

The distributions of stomach cancer mortality rates among nonwhites differed considerably from those among whites. Stomach cancer mortality rates for nonwhites among both sexes have been widespread and consistently high in southern Louisiana. Areas with moderate or low stomach cancer mortality rates were quite dispersed.

4.3 Factor Analysis of Cancer Mortality Rates in the U.S. by County

Factor analysis was undertaken to compress the selected cancer mortality rates into a more meaningful form to analyze the geographic patterns of cancer mortality in the U.S. counties from 1953 to 1977, from 1978 to 1987, and from 1953 to 1987. This section explains the results of factor analysis of cancer mortality rates in the U.S. from 1953 to 1987.

4.3.1 From 1953 to 1987

First of all, the appropriateness of the use of the factor model for this study was evaluated. The value of the test statistic for sphericity is large (=5057.12), and the associated significance level is small (significance = 0.0000), so it appeared unlikely that the population correlation matrix was an identity. The Kaiser-Meyer-Olkin (KMO) measure also showed a large value (0.60). These tests indicated that a factor analysis of the variables was appropriate for this study.

One of the estimates of the overall importance of any variable in a factor analysis is the communality value. Table 4.1 shows the communality of the variable which is the proportion of variance explained by the common factors. Lung cancer mortality rates for white males generated a communality of 0.91, the highest of all, and a uniqueness

Table 4.1 Communalities for Factor Analysis of Cancer Mortality Rates in the U.S. for 1953-1987

	Initial	Extraction
BNWF5387	.2539	.5598
BWF5387	.2665	.3367
CRNWF5387	.2427	.4177
CRNWM5387	.0315	.1045
CRWF5387	.2129	.2755
CRWM5387	.3861	.7073
LNWF5387	.0301	.2105
LNWM5387	.0997	.3282
LWF5387	.3176	.3439
LWM5387	.3615	.9129
PNWM5387	.0622	.0641
PWM5387	.0271	.0364
SNWF5387	.0230	.0593
SNWM5387	.0048	.0056
SWF5387	.1644	.4422
SWM5387	.1775	.2713

Extraction Method: Principal Axis Factoring

BNWF5387 - Nonwhite female breast cancer mortality from 1953 to 1987
BWF5387 - White female breast cancer mortality from 1953 to 1987
CRNWF5387 - Nonwhite female colorectal cancer mortality from 1953 to 1987
CRNWM5387 - Nonwhite male colorectal cancer mortality from 1953 to 1987
CRWF5387 - White female colorectal cancer mortality from 1953 to 1987
CRWM5387 - White male colorectal cancer mortality from 1953 to 1987
LNWF5387 - Nonwhite female lung cancer mortality from 1953 to 1987
LNWM5387 - Nonwhite male lung cancer mortality from 1953 to 1987
LWF5387 - White female lung cancer mortality from 1953 to 1987
LWM5387 - White male lung cancer mortality from 1953 to 1987
PNWM5387 - Nonwhite male prostate cancer mortality from 1953 to 1987
PWM5387 - White male prostate cancer mortality from 1953 to 1987
SNWF5387 - Nonwhite female stomach cancer mortality from 1953 to 1987
SNWM5387 - Nonwhite male stomach cancer mortality from 1953 to 1987
SWF5387 - White female stomach cancer mortality from 1953 to 1987
SWM5387 - White male stomach cancer mortality from 1953 to 1987

(Source: Calculated by author from cancer mortality rates of National Technical Information Service (NTIS), 1992)

(variance that is not explained by the common factors) of 0.09, which is the lowest of all. The high communality value for lung cancer for white males can be expected, given the dominance of this cancer type in many counties in the U.S. Colorectal cancer mortality rates for white males, which had second high communality (0.71), were concentrated in the Northeast. The lowest value was generated by stomach cancer mortality rates for nonwhite males (0.01). Such low value indicated little association of this cancer type with the others.

Table 4.2 contains statistics for each factor from the principal axis factoring extraction analysis. The total variance explained by each factor is listed in the column labeled Total. The next column contains the percentage of the total variance attributable to each resultant factor. The last column, the cumulative percentage, indicates the percentage of variance attributable to that factor and those that precede it in the table. In the initial statistics, six major factors were considered because only the factors with eigenvalue greater than 1 are meaningful. The first factor is the most important one in terms of capturing the variability of the entire set of variables and explains the largest amount of variance in the sample. The second factor accounts for the next largest amount of variance and is uncorrelated with the first. Successive factors explain extracted, the final six factors extracted explain 31.7 % of the total variance, as progressively smaller portions of the total sample variance, and all are uncorrelated with each other. The linear combination formed by factor 1 has a variance of 2.064, which has 12.9 % of the total variance. Factors 2, 3, 4, 5, and 6 explain 12.3%, 9.2%, 7.1%, 6.7%, and 6.5% of the variance, respectively. These six factors explain 54.7 % of the

Table 4.2 Total Variance Explained for Factor Analysis of Cancer Mortality Rates in the U.S. for 1953-1987

	Initial Eigenvalues			Extraction Sums of Squared Loadings			Rotation Sums of Squared Loadings		
Factor	Total	% of Variance	Cumulative %	Total	% of Variance	Cumulative %	Total	% of Variance	Cumulative %
1	2.064	12.901	12.901	1.549	9.678	9.678	1.335	8.346	8.346
2	1.967	12.292	25.193	1.495	9.344	19.023	1.264	7.902	16.248
3	1.467	9.170	34.363	.970	6.061	25.083	1.000	6.248	22.496
4	1.128	7.052	41.415	.509	3.180	28.264	.770	4.814	27.310
5	1.079	6.745	48.160	.354	2.210	30.474	.387	2.417	29.727
6	1.052	6.573	54.734	.200	1.252	31.725	.320	1.999	31.725
7	1.000	6.247	60.981						
8	.932	5.822	66.803						
9	.908	5.675	72.478						
10	.887	5.546	78.024						
11	.759	4.746	82.770						
12	.720	4.498	87.269						
13	.691	4.320	91.588						
14	.517	3.232	94.820						
15	.435	2.716	97.536						
16	.394	2.464	100.000						

Extraction Method: Principal Axis Factoring

(Source: Calculated by author from NTIS cancer mortality rates, 1992)

total variance of the cancer mortality rates. The remaining 10 factors together account for only 45.3 % of the variance. This model, incorporating the six significant factors, was adequate to represent the data. After the desired number of factors has been extracted, the final six factors extracted explain 31.7 % of the total variance, as compared to 54.7 % for the first six factors in the initial statistics. However, the eigenvalues on factors 5 and 6 from extraction and rotation of sums of squared loadings were too small, and the factors might not be meaningful.

The Scree test was reproduced in Figure 4.12. It is a graph of the initial eigenvalues plotted on the factors. It plots the proportion of total variance associated with each factor. The plot shows a distinct break between the steep slope of the large factors and the gradual trailing off of the rest of the factors. After four factors, the plot becomes essentially flat. This reflects the importance of the first four factors.

Rotated factor matrix on Table 4.3 shows the coefficients used to express each standardized variable in terms of the factors. These coefficients or factor loadings indicate the contribution of each variable to the explanatory power of each factor. Large factor loadings for a variable indicate that this factor is closely related to the variable. Here, the rotated varimax method was used to make the factor loadings more interpretable. Its purpose is to achieve a simple structure. Although the factor matrix changes, the communalities and the percentage of total variance explained do not change. Therefore, both the rotated and the unrotated factors reproduce the correlations with the same degree of accuracy. This transformation merely simplifies the discussion of the contribution of each variable on each factor.

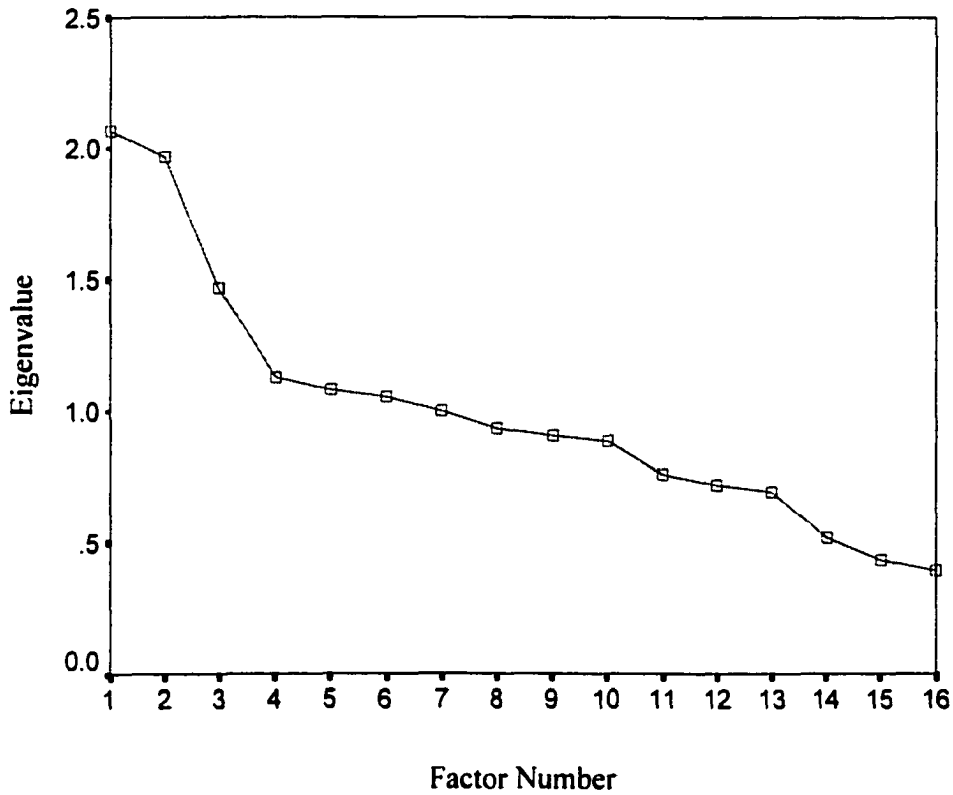

Figure 4.12 Factors and Their Eigenvalues in the U.S. for 1953-1987

Table 4.3 Rotated Factor Matrix in the U.S. for 1953-1987

	Factor 1	Factor 2	Factor 3	Factor 4	Factor 5	Factor6
LWM5387	.9440	-.0816	.0124	-.0532	.0365	.1036
LWF5387	.5559	.1485	-.0020	-.0619	.0870	.0370
PNWM5387	.2207	.0232	.0753	-.0921	.0225	.0152
SNWM5387	.0573	-.0182	-.0043	-.0062	.0025	-.0442
CRWM5387	.1058	.7840	.0439	.2714	.0531	.0563
BWF5387	-.0016	.5609	.0076	.1068	.0969	.0355
CRWF5387	-.0130	.5183	.0293	.0159	.0414	.0623
BNWF5387	.0400	.0188	.7374	-.0105	.0635	.1002
CRNWF5387	.0428	.0409	.6390	-.0053	.0456	.0614
SWF5387	-.1331	.0046	-.0001	.6508	.0026	.0295
SWM5387	-.1892	.1550	.0139	.4587	.0037	-.0300
PWM5387	.0087	.0558	-.0113	.1800	-.0104	-.0234
LNWM5387	.1113	.0288	.1792	-.0482	.4988	.1783
CRNWM5387	.0211	.0711	-.0149	.0083	.3133	-.0240
LNWF5387	-.0162	.0358	-.0059	-.0230	.0991	.4457
SNWF5387	.0400	.0245	.0793	-.0188	-.0181	.2238

Extraction Method: Principal Axis Factoring
Rotation Method: Varimax

See Table 4.1 for footnotes.
(Source: Calculated by author from NTIS cancer mortality rates, 1992)

In the first factor, the heaviest loadings were from cancers of white male lung (0.94) and white female lung (0.56), indicating similar spatial patterns among these two cancers. The first factor showed a strongly positive correlation with lung cancer for white males. The heaviest loading on the second factor was of colorectal cancer for white males (0.78). Cancers of breast (0.56) and colorectum (0.52) for white females also emerged with heavier loadings on the second factor. The heaviest loadings on the third factor were cancers of breast (0.74) and colorectum (0.64) for nonwhite females, and this factor was highly and positively correlated with these cancers. The fourth factor loaded heaviest on stomach cancer for white females (0.65) and heavier on

the stomach cancer for white males (0.46). This factor showed a relatively high correlation with stomach cancer for whites. Factors 5 and 6 loaded heaviest on lung cancers for nonwhite males (0.50) and for nonwhite females (0.45), respectively. But the last two factors were not considered in this analysis because these two factors showed low factor loadings below 0.5 (Table 4.3) and had eigenvalue less than 1 (the total variance explained on rotation sums of squared loadings in Table 4.2).

Factor scores of the first four factors explained above were demonstrated in maps of U.S. counties (Figures 4.13 - 4.16). Factor scores of U.S. counties were depicted in terms of five classes determined according to the values: i.e. (a) less than -1.5 (lower), (b) -1.5 ~ -0.5 (low), (c) -0.5 ~ 0.5 (modest), (d) 0.5 ~ 1.5 (high), and (e) more than 1.5 (higher). Counties with the highest scores of factor 1 (Figure 4.13) were Loving, Texas (4.30), and Charlton, Georgia (3.84), Reagan, Texas (3.19), Crisp (3.19) and McIntosh (3.14) in Georgia, and St. Bernard, Louisiana (3.13). Exceptionally, higher scores were evident in most of the southeastern U.S., from South Georgia through North Florida to the southern counties of Alabama, Mississippi, and Louisiana. Relatively high scores generated in some counties of the northeastern U.S, Pacific regions, and the South. Far more modest and low scores prevailed in most of the Midwest and North Mountain regions. Lower scores were located predominantly in much of the South Mountain region. Counties with the lowest scores were Daggett, Utah (-3.66) and Conejos, Colorado (-3.06), respectively.

In the second factor (Figure 4.14), the highest scores were in Shannon, South Dakota (9.22), Camas, Idaho (3.24), and Glasscock, Texas (3.10) and the lowest scores were in Kenedy, Texas (-3.59), and Chattahoo (-2.62) and Clay (-2.46), Georgia. Higher

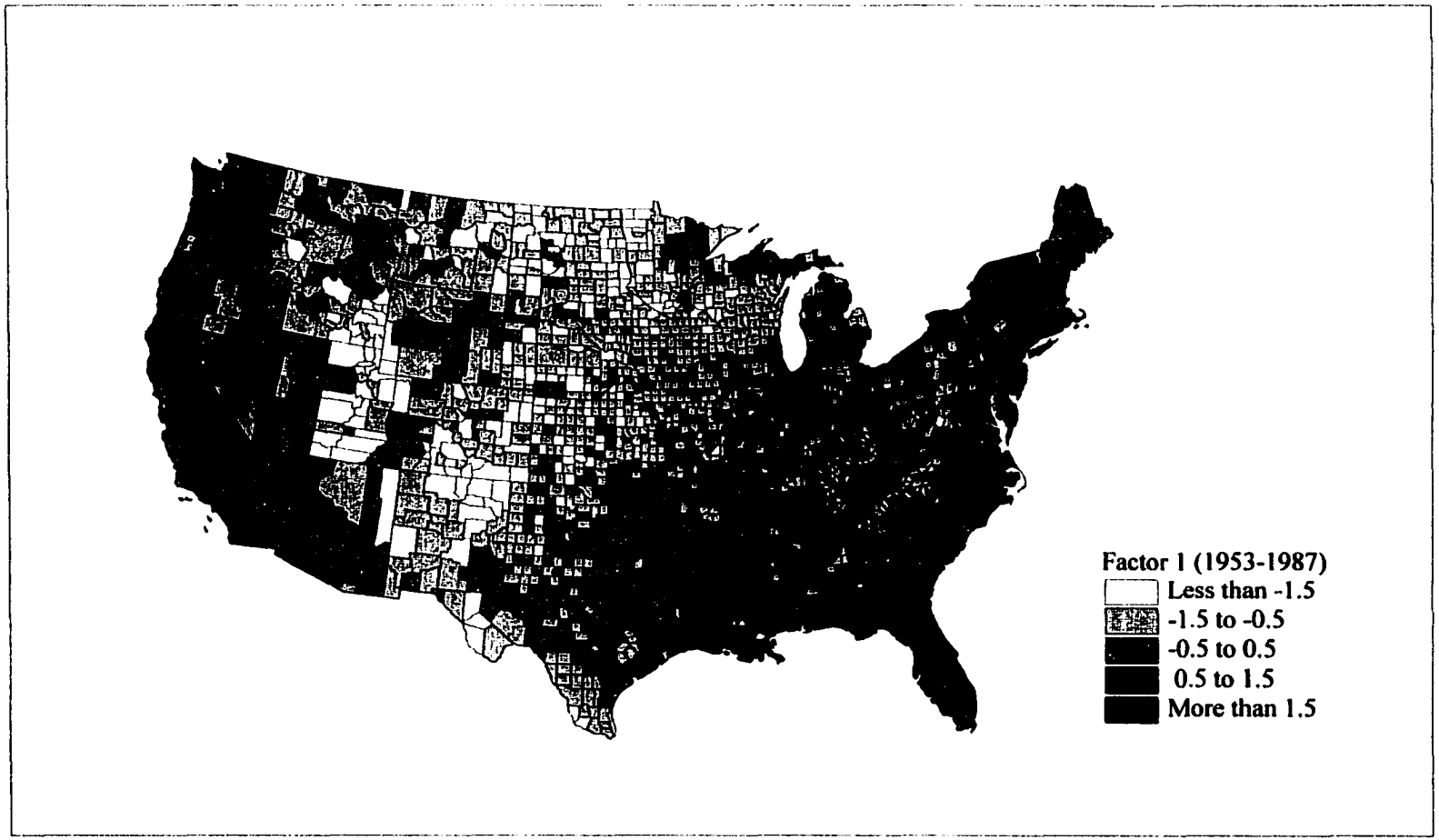

Figure 4.13 Factor Score Distribution for Factor 1, U.S. counties (1953-1987)
(Source: Calculated by Author from National Technical Information Service, 1992)

Figure 4.14 Factor Score Distribution for Factor 2, U.S. counties (1953-1987)
(Source: Calculated by Author from National Technical Information Service, 1992)

scores prevailed in some counties of the Northeast. High scores were more dispersed and generated by most of the Northeast and East North Central regions. Most Pacific regions, West North Central regions, and Florida developed moderate to slightly low scores. Most of the South (except for Florida) showed relatively lower scores.

A wide range of scores developed on the third factor but most places generated a value between 0.5 and -0.5 (Figure 4.15). Counties with the highest scores were Elk, Pennsylvania (34.75), Clay, Kansas (14.74), Tucker, West Virginia (6.81), Garfield, Montana (6.76), Pendleton, West Virginia (5.69), and Crook, Wyoming (5.43). The highest scores of factor 3 were higher than those of other factors, and their distributions were more dispersed. The lowest scores were in Kalkaska, Michigan (-1.10), Floyd, Iowa (-0.96), and Douglas, Minnesota (-0.92). The lowest scores of factor 3 were much higher than those of others factors. The distributions of high and low scores for this factor were dispersed but their ranges were relatively narrow. Most counties of the U.S. developed moderate scores, but positive factor scores (0 ~ 0.5) were usually generated in the eastern half and negative factor scores (-0.5 ~ 0) in the western half of the U.S.

With the fourth factor (Figure 4.16), the highest scores were in Loving, Texas (18.13); Alpine, California (6.90); and Summit, Colorado (5.31). Most of northern Midwest regions and New Mexico, and some of the northern U.S. showed relatively higher scores. Most counties of the U.S developed moderate to low scores, with the exception of Nevada, New Mexico, northern Midwest regions, and the northeastern states. Counties with the lowest scores were in King, Texas (-3.17) and Shannon, South Dakota (-3.39).

Figure 4.15 Factor Score Distribution for Factor 3, U.S. counties (1953-1987)
(Source: Calculated by Author from National Technical Information Service, 1992)

Figure 4.16 Factor Score Distribution for Factor 4, U.S. counties (1953-1987)
(Source: Calculated by Author from National Technical Information Service, 1992)

The following section discusses in detail four major factors extracted.

1. Cancers of white male lung (0.94) and white female lung (0.56) loaded highly onto factor 1. This factor was highly correlated with lung cancer of white male and exhibited positive correlations of the remaining cancers, except for stomach cancers for white males (-0.19) and females (-0.13), lung cancer for nonwhite females (-0.02), colorectal cancer for white females (-0.01), and breast cancer for white females (0.002). Counties which had high scores on this factor had higher rates than average mortality rates from lung cancers for white males and females. Counties which had low scores on this factor showed lower than average mortality rates from these lung cancers for whites. High scores for this factor were located predominantly in the southeastern U.S., parts of Pacific regions, and Northeast. Low scores prevailed in most of the Midwest and Mountain regions. The highest loading variable (cancer) in each factor was chosen to represent the information conveyed in that factor. Major cancers representing factor 1 were related to lung cancers in whites.

Even though this study did not examine the correlation coefficients between factor scores and some environmental variables, previous studies showed the association between environmental factors and lung cancer (Blot et al. 1976, 1979, and 1982; Greenberg 1983; Hoover and Fraumeni 1975; Howe 1981; Page et al. 1985). Lung cancer was the most frequently diagnosed cancer and the leading cause of cancer death among Americans from 1953 to 1987. Compared with other common cancers such as breast, colorectum or prostate, lung cancer displayed much greater geographic variation. The patterns of lung cancer mortality rates identified in this study were consistent with the incidence statistics. High rates of lung cancer mortality in Atlantic Coast areas and

Louisiana have been associated with employment in chemicals, petroleum, transportation, and paper industries (Blot and Fraumeni 1976). Another study suggested that the high cancer death rates observed for Louisiana were probably not due to high incidence but a result of poor prognosis from later or more advanced tumor stage at the time of diagnosis (Chen et al. 1997). The analysis of cancer mortality in the British Isles identified the urban affinity of lung cancer (Howe 1981). Numerous environmental exposures have been implicated as etiologic risk factors for lung cancer, including air pollution, indoor radon, tobacco smoke, asbestos, ionizing radiation, and a number of occupational agents.

In particular, lung cancer was identified as a member of the 'smoking' association (Hoover and Fraumeni 1975). Smoking accounts for an overwhelming share of lung cancer deaths in the U.S., increasing an individual's risk of dying from lung cancer by tenfold compared with the nonsmoker (Aronchick 1990). Unfortunately, historical tobacco use data were not available. But Current Population Survey by U.S. Census Bureau in 1985 showed that the state-by-state pattern of lung cancer mortality rates is very similar to the pattern of smoking adult rates (U.S. Census Bureau 1985). In the 1950s, seven of the eight states with the lowest cigarette consumption were in the South, and these southern states had below-average rates of lung cancer (Greenberg 1984). In the 1990s, the South showed seven of the eight states with the highest lung cancer mortality rates, but only three of the eight sates (Kentucky, West Virginia, and Oklahoma) had the largest percentages of their populations who smoke (U.S. Department of Health and Human Services 1989). Why the southern states show up more prominently for lung cancer mortality than for smoking may be due to a number of

factors. The average southern smoker may smoke more heavily than the average northern smoker (Goldman 1991). In case of tax-paid per capita sales in number of packs, the South has shown below-average (or average) tax-paid per capita sales in the 1950s to the 1990s, but they had above-average tax-paid per capita sales in the 1990s (Tobacco Institute 1994). Therefore, the first factor could be interpreted as lung cancers for whites and the factor regions (that have the highest factor score) were identified as the South.

2. Colorectal cancer for white males (0.78), breast cancer for white females (0.56), and colorectal cancer for white females (0.52) loaded highly onto factor 2. The second resultant factor was strongly and positively correlated with colorectal cancer for white males. This factor showed positive correlations of the remaining cancers, except for lung cancer for white males (-0.08) and stomach cancer for nonwhite males (-0.02). Counties which had high scores on factor 2 had higher than average mortality rates from colorectal cancer for white males. Counties which had low scores on this factor showed lower than average cancer mortality rates from colorectal cancer for white males. The Northeast and East North Central regions showed high scores on factor 2 and most of the South (except for Florida) showed low scores. Major cancers representing factor 2 were related to colorectal cancers for whites and breast cancer for white females. This factor regions were identified as the northeastern U.S.

Using 52 nations as a base line, an association was found between urbanization and the following types of cancer: male and female colon, male rectum, male and female lung. In particular, colorectal cancers occurred more often in urban, industrialized countries than in rural countries, but Japan was an exception (Page and Asire 1985). The

most urban counties in the U.S. had white population cancer rates at least 40 percent higher than the least urban counties for cancers of male and female colorectum, female breast, and male lung from 1950 to 1975 (Greenberg 1983).

In the 1980s, the geographic pattern of state population densities still showed some major features which have been consistent since 1900. The most densely populated counties were still those of the Northeast. These states also had a large percentage of European-born, white Americans, relatively high socioeconomic status in comparison to the rural states, and a large concentration of production workers. There was a striking concentration of the highest cancer rates in the Northeast, especially in urban and suburban areas, and this has been the case for many years. The highest concentration of breast cancer mortality rates has been located primarily in the Northeast. Women with higher socioeconomic status have been more likely to get breast cancer (Goldman 1991). White Americans have relatively high socioeconomic status in comparison to nonwhites. Therefore, this factor located predominantly in the northeastern U.S. could be interpreted as the cancers for whites related to urbanization.

3. Breast cancer for nonwhite females (0.74) and colorectal cancer for nonwhite females (0.64) loaded highly onto factor 3, indicating a strong relationship between this factor and the two cancers. Counties which had high scores on this factor had higher rates than average cancer mortality rates from nonwhite female breast and nonwhite female colorectum. Counties which had low scores on this factor tended to have lower than average cancer mortality rates from these. It was difficult to interpret the factor and nominate its regions because the factor scores were relatively moderate and widely

dispersed. But positive factor scores were usually shown in the eastern half and negative factor scores in the western half of the U.S.

Just as the population densities of areas vary, so too do their racial compositions. The proportions of black, white, and other population groups have remained stable throughout the 20th century. But there were substantial differences from area to area in terms of the racial characteristics of the population. Most of the Indians and Orientals that are part of the nonwhite group were in the western half of the country whereas the blacks were clustered in the eastern half. The proportions of blacks in the populations of most northern and western states have increased, but the states in the southeastern part of the country continue to have the largest proportions of blacks. These nonwhites have relatively low socioeconomic status in comparison to white Americans. Before World War II, nonwhites had lower rates of cancer than whites. Nonwhites (African Americans in particular) have had the highest risk of cancers and many other diseases since World War II (Goldman 1991). Factor 3 might be interpreted as the cancers related to urbanization of nonwhite females.

4. Stomach cancers for white females (0.65) and males (0.46) loaded relatively highly onto factor 4. This factor had a highly positive correlation with stomach cancer for white females. Counties which had high scores on factor 4 had higher than average mortality rates from stomach cancer for white females. Counties which had low scores on this factor tended to have lower than average cancer mortality rates from stomach cancer for white females. Most of the northern Midwest regions, some of the northern states and New Mexico showed relatively high factor scores.

International survey had shown that some of the most urbanized nations (Australia, New Zealand, Canada, and the U.S.) had low stomach cancer mortality rates, except for Japan. Some of the least urbanized nations (Bulgaria, Costa Rica, Romania, and Ecuador) had high stomach cancer mortality rates. According to a study of the relationship between urbanization and cancer mortality rates in the U.S., cancers of the stomach and prostate were of higher rates in the rural than in the urban counties (Greenberg 1983). Low socioeconomic status was associated with stomach cancer (Normura 1982).

The cancers of stomach and colorectum are cancers of digestive systems. Many cases of male and female stomach cancer mortality rates could be more directly attributed to behavioral and dietary customs associated with some ethnic groups. For example, a concentration in certain southwestern states seemed to be related to the excess risk among the Hispanic groups and American Indians in this area (Wiggins et al. 1989). Therefore, major cancers identifying factor 4 were related to stomach cancers of the digestive system for all the whites. The concentrations of high factor score areas were identified as New Mexico and the upper West North Central region.

4.4 Summary

This chapter began with three approaches. The first and second are an examination of the geographic patterns of cancer mortality rates from the contiguous U.S. states for the period 1953-1987 at the state and county level, respectively. The third is a factor analysis for the five cancer mortality rates in the U.S. counties from 1953 to 1987. The results are summarized as follows:

1. The northeastern U.S. during 1953-1987 showed the highest cancer mortality rates. High cancer regions were along the West coast (such as Washington, Nevada, and California) and South coast of the U.S. (such as Louisiana and Florida). The Great Plains and Rocky Mountain areas generally generated low cancer death rates than the national average. Males had higher cancer mortality rates than females of the same race. Cancer occurred more frequently among nonwhites. The spatial patterns between males and females for the same cancer were generally very similar whereas those between whites and nonwhites were different.

2. The geographic distributions of cancer mortality rates at the county level were relatively similar to those at the state scale. The mortality patterns for most cancers showed a general tendency toward geographic uniformity over time. Of the 5 cancer sites examined at the county level, the distributions in breast and colorectal cancer mortality rates developed similar spatial patterns; high rate areas were markedly concentrated in the northeastern quadrant and low rates were shown in the southern half of the country. The areas of high lung cancer mortality rates prevailed in the southeastern States, the West coast, and in the northeastern part of the country, whereas those of low rates were dispersed in the Midwest and Mountain region. Distributions of high mortality rates for prostate cancer generated in counties of South Carolina, North Carolina, and Utah, whereas those of low rates occurred in most of southern West regions. For stomach cancer mortality, areas of high rates developed in the upper Midwest and northern New Mexico, and southern Louisiana, but not in the Northeast.

3. In factor analysis for the 16 cancer types of 1953-1987, four factors were extracted and examined. Major cancers representing factor 1 were related to lung

cancers for whites. The first factor could be interpreted as "lung cancers for whites" and the factor regions were identified as the South. The second factor was strongly and positively correlated with the cancers of colorectum and breast for whites. The second could be defined as the "cancers of colorectum and breast for whites." This factor's regions were concentrated in the northeastern States. Major cancers identifying factor 3 were related to the cancers of colorectum and breast for all the nonwhite females. It was not easy to interpret the factor and identify its regions because the factor scores were widely dispersed. But positive factor scores were usually located in the eastern half of the country. Factor 3 might be interpreted as the "cancers of colorectum and breast for nonwhite females." The fourth factor was highly and positively related to stomach cancers among whites and could be considered as "cancers of digestive system for whites." The concentrations of high factor score areas were identified as New Mexico and the upper West North Central region.

As unit areas for analysis, state and counties in the U.S. are the most convenient, although other subdivisions may be used where data are available. Counties are preferable to states in terms of sensitivity to geographic variation in mortality, but they are prone to large errors where rates are based on small population and on death numbers drawn from short time periods. Given the disparities of cancer mortality rates from 1953 to 1987, the advantages of the county scale were evident. Furthermore, it is important to recognize that some geographic and temporal variations for mortality rates provide signals to environmental hazards, but for many tumors fluctuations in medical care, diagnosis, reporting, survival time, and migration may complicate the picture. Also, many of the coastal counties in Louisiana have manifested consistently high rates and

developed distinctive spatial patterns for lung cancer mortality rates. This analysis revealed the basis of the first hypothesis that cancer mortality rates are higher in South Louisiana than the nation, which is discussed in Chapter 5 in detail.

CHAPTER 5

LOUISIANA

This chapter analyzes the geographic patterns of cancer mortality rates within Louisiana parishes from the 1950s to the 1980s in detail. It also examines the three hypotheses: Cancer mortality rates are higher in South Louisiana than the nation and the state; there are spatial clusterings of cancer mortality rates of some major sites including breast, colon and rectum, lung, prostate, and stomach; and these cancer mortality occurrences are associated with environmental variables. For these hypotheses, significance test of rate differences, factor analysis, spatial autocorrelation, and multiple regression are employed and the results are explained in each section.

5.1 Comparison of Cancer Mortality Rates in South Louisiana and the U.S.: Significance Test

Age-adjusted mortality rates for each of the 35 South Louisiana parishes and for the entire state were compared with the U.S. rates. Standard errors of the age-adjusted rates of each parish were obtained from NTIS and their rate ratios were calculated in this study so that the significance of rate differences could be assessed (NTIS 1992, Esteve et al. 1994). For example, statistical significance is based upon calculating the probability that any difference between a parish rate and the national rate is due to chance alone. This statistical significance is a function of both the magnitude of the difference between the parish and national rate and the population size of the parish. The cutoff value for the significance test statistic is $p < 0.05$ (NTIS 1992). This test is relatively stable and reliable since common cancers have been computed from large numbers of cases.

All Sites Combined: Table 5.1 shows that Orleans Parish had the state's highest mortality rates (202.35) for all cancers combined, whereas the lowest rates were observed in East Feliciana during 1953-1987. In Table 5.1, more than half of the parishes in South Louisiana (19 out of 35 parishes) had significantly higher cancer mortality than the U.S.

Only five parishes had significantly lower mortality rates than the U.S., and the rest of them had similar rates to the U.S rates (Figure 5.1). Age-adjusted mortality rates for all cancers combined showed significantly higher rates for males than females, and for nonwhites than whites. Rates of all sites combined by sex and race were summarized in Appendices I and J.

Parishes which showed significantly high cancer mortality rates for white males were very similar to those for all sites combined and were more commonly found in South Louisiana than those for nonwhite males. In the case of white females, only Orleans Parish had significantly higher rates than the U.S. Cancer mortality rates among nonwhite females were higher than or similar to those of the U.S, except for East Feliciana, St. Helena, Tangipahoa, and Washington parishes.

Lung Cancer: In Louisiana, lung cancer is the leading cause of cancer death (average annual age-adjusted rates were 40.33 deaths per 100,000). St. Bernard (53.43) and St. Tammany (51.23) had the highest rates, whereas East Feliciana (19.35) and Claiborne (25.84) had the lowest rates (Table 5.1). As depicted in Figure 5.2 and shown in Table 5.1, most parishes (28) in South Louisiana had significantly higher cancer mortality rates for lung cancer (both sexes and races) than the U.S. Only East Feliciana

Table 5.1 Age-adjusted Cancer Mortality Rates: Significance Test Results, 1953-1987

	All Sites: Both Sexes and Races			Lung: Both Sexes and Races		
South Louisiana	Rate	S.E.	Rate Ratio	Rate	S.E.	Rate Ratio
Acadia	188.08	6.53	1.15**	47.79	3.28	1.41**
Allen	166.05	9.41	1.01	36.94	4.44	1.09
Ascension	165.61	7.84	1.01	40.01	3.86	1.18**
Assumption	173.00	10.53	1.05	42.12	5.21	1.25**
Beauregard	162.34	8.80	0.99	37.38	4.24	1.10
Calcasieu	175.14	4.13	1.07**	45.95	2.10	1.36**
Cameron	143.38	15.29	0.87*	45.92	8.58	1.36**
East Baton Rouge	164.20	2.99	1.00	36.94	1.41	1.09**
East Feliciana	106.68	7.33	0.65*	19.35	3.13	0.57*
Evangeline	182.06	8.15	1.11**	48.44	4.16	1.43**
Iberia	191.56	6.75	1.17**	47.95	3.37	1.42**
Iberville	165.65	7.99	1.01	36.81	3.77	1.09
Jefferson	178.11	2.97	1.09**	48.54	1.53	1.43**
Jefferson Davis	180.19	8.44	1.10**	42.07	4.07	1.24**
Lafayette	177.46	5.02	1.08**	43.94	2.49	1.30**
Lafourche	169.52	6.25	1.03	39.97	3.02	1.18**
Livingston	150.91	7.42	0.92*	43.69	3.97	1.29**
Orleans/N`Orlean	202.35	1.94	1.23**	44.38	0.90	1.31**
Plaquemines	189.05	12.24	1.15**	46.85	5.87	1.38**
Pointe Coupee	168.12	9.14	1.02	34.42	4.13	1.02
St. Bernard	187.80	7.97	1.14**	53.43	4.19	1.58**
St. Charles	164.87	9.70	1.00	46.50	5.12	1.37**
St. Helena	113.78	11.81	0.69*	27.80	5.78	0.82*
St. James	178.35	11.02	1.09**	39.25	5.21	1.16**
St. John Baptist	185.44	10.78	1.13**	42.79	5.16	1.26**
St. Landry	178.31	5.27	1.09**	42.05	2.55	1.24**
St. Martin	188.45	8.81	1.15**	46.99	4.38	1.39**
St. Mary	186.98	7.10	1.14**	45.36	3.49	1.34**
St. Tammany	172.40	5.64	1.05**	51.23	3.06	1.51**
Tangipahoa	162.94	5.38	0.99	36.45	2.54	1.08**
Terrebonne	182.03	6.46	1.11**	47.30	3.27	1.40**
Vermilion	175.51	6.68	1.07**	37.91	3.10	1.12**
Washington	157.36	6.34	0.96*	38.00	3.10	1.12**
West Baton Rouge	179.29	12.22	1.09**	40.89	5.82	1.21**
West Feliciana	147.69	14.46	0.90*	31.56	6.64	0.93
LA	172.96	0.77	1.05**	40.33	0.37	1.19**
U.S.	164.13	0.09	1.00	33.83	0.04	1.00

Deaths per 100,000, adjusted to the age distribution of the 1970 U.S. Population
S.E.: Standard error of the age-adjusted rate
Rate Ratio: Ratio of each parish rate to U.S. rate
** (*): Significantly higher (lower) than the U.S. rates at the 0.05 level
(Source: calculated by author from NTIS, 1992)

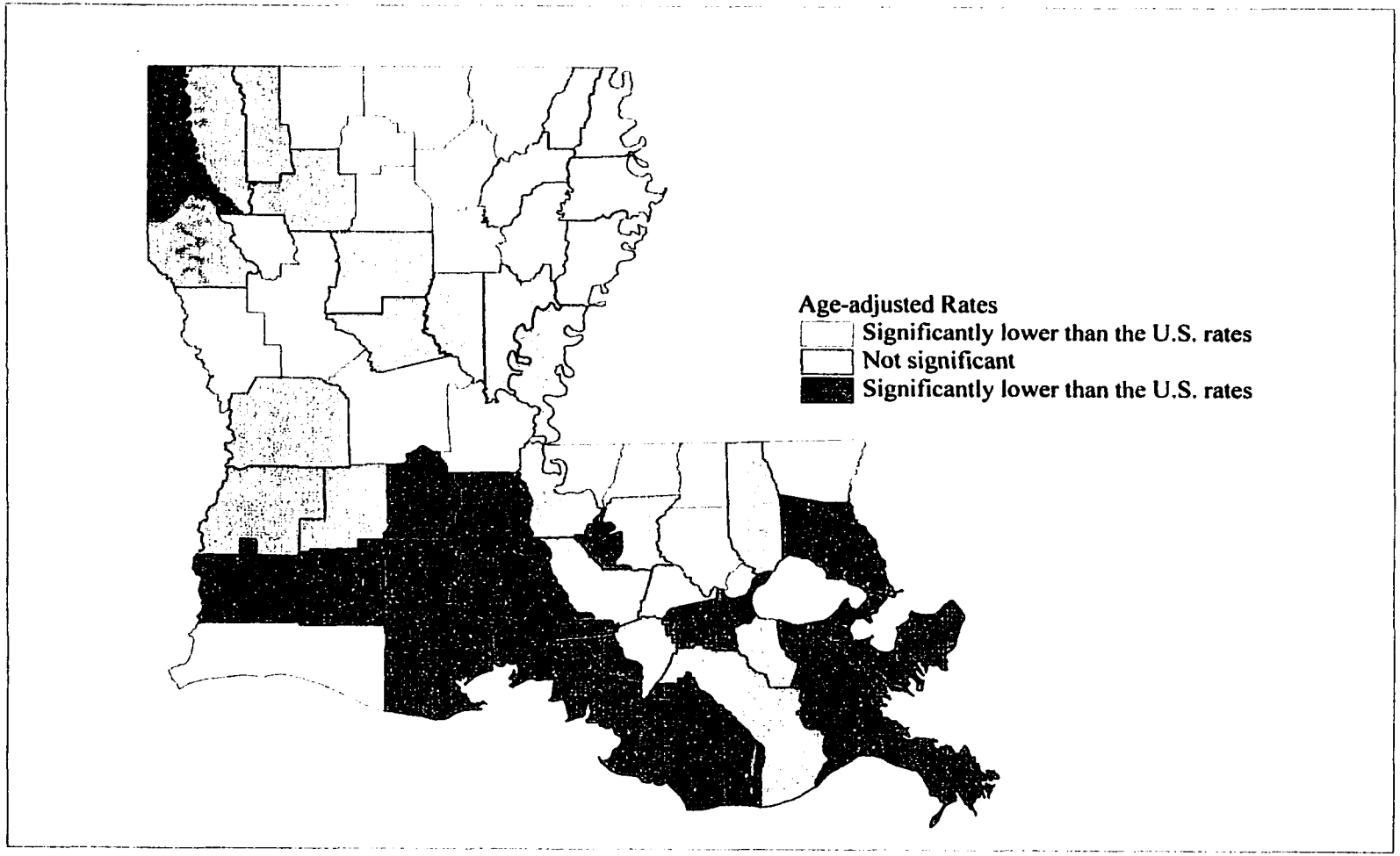

Figure 5.1 Cancer Mortality Rates (All sites combined) Louisiana: 1953-1987
(Source: Modified by Author from National Technical Information Service, 1992)

Figure 5.2 Lung Cancer Mortality Rates, Louisiana: 1953-1987
(Source: Modified by Author from National Technical Information Service, 1992)

and St. Helena showed significantly lower rates of lung cancer, and five parishes had similar mortality rates to the U.S rates.

White males in Louisiana had the highest cancer mortality of lung cancer (73.77) (Table 5.2). Most parishes (26) in South Louisiana had significantly higher cancer mortality rates and only East Feliciana had significantly lower rates than the U.S. (Table 5.2). In the case of lung cancer mortality rates among white females, Assumption, East Feliciana, Lafourche, and St. James parishes indicated significantly lower rates than the U.S., and other parishes showed higher rates than, or similar rates to, the U.S. rates (Appendix K). Lung cancer mortality rates for nonwhites were significantly higher or similar to those of the U.S., except for East Feliciana, St. Helena, and Washington parishes (Appendix L).

Stomach Cancer: Twenty-one of 35 parishes in South Louisiana showed significantly higher cancer mortality rates for stomach cancers (both sexes and races) than the U.S. (Table 5.3 and Figure 5.3). The parishes that showed significantly high rates had big difference among sexes and races. Stomach cancer mortality rates for males and nonwhites were two times higher than those females and whites, respectively (Appendices M - N). Cancer mortality rates among white males for South Louisiana were either similar to, or lower than those for the U.S., except for St. Landry and St. Martin parishes (Appendix M). The comparisons of stomach cancer among white females did not show significantly high rates in South Louisiana (Appendix N). On the contrary, stomach cancer among nonwhite females did not show any significantly lower rates than that of the U.S. (Appendix N). In particular, more than half of the parishes in South Louisiana had significantly higher cancer mortality rates for nonwhite

Table 5.2 Age-adjusted Cancer Mortality Rates: Significance Test Results, 1953-1987

	Lung: White Males			Lung: Nonwhite Males		
South Louisiana	Rate	S.E.	Rate Ratio	Rate	S.E.	Rate Ratio
Acadia	83.43	7.11	1.47**	70.90	14.97	1.09
Allen	67.57	9.91	1.19**	59.20	18.15	0.91
Ascension	77.09	9.43	1.35**	56.92	13.51	0.88
Assumption	81.90	13.10	1.44**	79.89	19.10	1.23
Beauregard	60.58	8.53	1.06	70.81	21.80	1.09
Calcasieu	78.61	4.66	1.38**	85.41	9.53	1.32**
Cameron	68.24	15.56	1.20	96.46	80.48	1.49
East Baton Rouge	67.39	3.64	1.18**	61.46	5.03	0.94
East Feliciana	36.72	8.50	0.65*	32.53	8.96	0.50*
Evangeline	82.35	8.82	1.45**	78.97	19.77	1.22
Iberia	85.84	7.98	1.51**	83.91	12.95	1.30**
Iberville	65.84	10.03	1.16	58.20	10.74	0.90
Jefferson	84.59	3.41	1.49**	97.63	9.52	1.51**
Jefferson Davis	72.75	8.79	1.28**	65.58	18.04	1.01
Lafayette	75.04	5.66	1.32**	89.30	12.49	1.38**
Lafourche	75.02	6.66	1.32**	69.78	18.29	1.08
Livingston	76.49	8.10	1.34**	70.86	24.71	1.09
Orleans/N`Orlean	79.01	2.34	1.39**	89.18	3.34	1.38**
Plaquemines	73.48	12.38	1.29**	79.40	21.22	1.23
Pointe Coupee	58.08	10.31	1.02	58.74	12.27	0.91
St Bernard	96.30	9.28	1.69**	88.09	34.15	1.36
St Charles	81.06	12.18	1.42**	95.86	21.56	1.48**
St Helena	64.89	18.22	1.14	32.27	13.31	0.50
St James	75.76	14.92	1.33**	79.22	16.91	1.22
St John Baptist	66.33	13.08	1.17	85.19	16.58	1.32**
St Landry	73.50	6.33	1.29**	68.80	8.23	1.06
St Martin	82.62	10.27	1.45**	85.14	16.32	1.31**
St Mary	79.83	8.54	1.40**	78.47	12.28	1.21**
St Tammany	84.70	6.38	1.49**	85.09	14.79	1.31**
Tangipahoa	67.79	6.09	1.19**	57.44	9.23	0.89
Terrebonne	89.07	7.61	1.57**	81.45	14.64	1.26**
Vermilion	67.41	6.56	1.18**	69.59	20.15	1.07
Washington	76.01	7.76	1.34**	52.76	10.50	0.81*
West Baton Rouge	74.31	16.13	1.31**	76.38	17.76	1.18
West Feliciana	42.41	17.37	0.75	51.88	15.84	0.80
LA	73.77	0.89	1.30**	66.94	1.36	1.03**
U.S.	56.90	0.09	1.00	64.76	0.29	1.00

See Table 5.1 for footnotes.
(Source: calculated by author from NTIS, 1992)

Table 5.3 Age-adjusted Cancer Mortality Rates: Significance Test Results, 1953-1987

	Stomach: Both Sexes and Races			Stomach: Nonwhite Males		
South Louisiana	Rate	S.E.	Rate Ratio	Rate	S.E.	Rate Ratio
Acadia	10.12	1.52	1.22**	22.83	8.50	1.21
Allen	10.00	2.30	1.21	22.68	11.14	1.21
Ascension	11.46	2.07	1.38**	26.40	9.20	1.40
Assumption	12.25	2.82	1.48**	41.07	13.68	2.18**
Beauregard	7.90	1.94	0.95	19.78	11.25	1.05
Calcasieu	8.56	0.93	1.03	29.25	5.64	1.55**
Cameron	5.26	2.91	0.63*	17.70	34.70	0.94
East Baton Rouge	7.70	0.66	0.93	20.92	3.01	1.11
East Feliciana	7.20	1.89	0.87	15.92	6.27	0.85
Evangeline	9.49	1.85	1.14	32.20	12.50	1.71**
Iberia	13.05	1.77	1.57**	44.49	9.59	2.36**
Iberville	12.56	2.20	1.52**	29.85	7.58	1.59**
Jefferson	7.16	0.61	0.86*	22.86	4.61	1.21
Jefferson Davis	10.71	2.07	1.29**	32.87	12.93	1.75**
Lafayette	9.87	1.20	1.19**	29.17	7.09	1.55**
Lafourche	10.66	1.59	1.29**	41.89	14.28	2.23**
Livingston	4.76	1.35	0.57*	19.56	12.94	1.04
Orleans/N`Orlean	10.99	0.45	1.33**	27.23	1.89	1.45**
Plaquemines	11.48	3.02	1.38**	29.51	12.69	1.57
Pointe Coupee	12.70	2.51	1.53**	24.84	7.94	1.32
St Bernard	7.42	1.61	0.90	54.61	29.97	2.90**
St Charles	11.10	2.56	1.34**	26.35	11.35	1.40
St Helena	6.88	2.89	0.83	14.41	9.03	0.77
St James	15.90	3.29	1.92**	40.68	12.23	2.16**
St John Baptist	14.63	3.06	1.76**	32.45	10.28	1.72**
St Landry	13.77	1.47	1.66**	29.71	5.38	1.58**
St Martin	13.55	2.37	1.63**	28.58	9.58	1.52**
St Mary	11.50	1.77	1.39**	33.09	7.90	1.76**
St Tammany	6.50	1.10	0.78*	15.97	6.43	0.85
Tangipahoa	10.46	1.37	1.26**	29.27	6.64	1.56**
Terrebonne	11.23	1.62	1.35**	37.73	10.21	2.00**
Vermilion	10.18	1.61	1.23**	33.58	14.35	1.78**
Washington	7.42	1.37	0.90	18.28	6.18	0.97
West Baton Rouge	13.33	3.37	1.61**	22.49	9.47	1.20
West Feliciana	11.16	4.00	1.35	15.87	8.76	0.84
LA	9.36	0.18	1.13**	24.22	0.82	1.29**
U.S.	8.29	0.02	1.00	18.82	0.16	1.00

See Table 5.1 for footnotes.
(Source: calculated by author from NTIS, 1992)

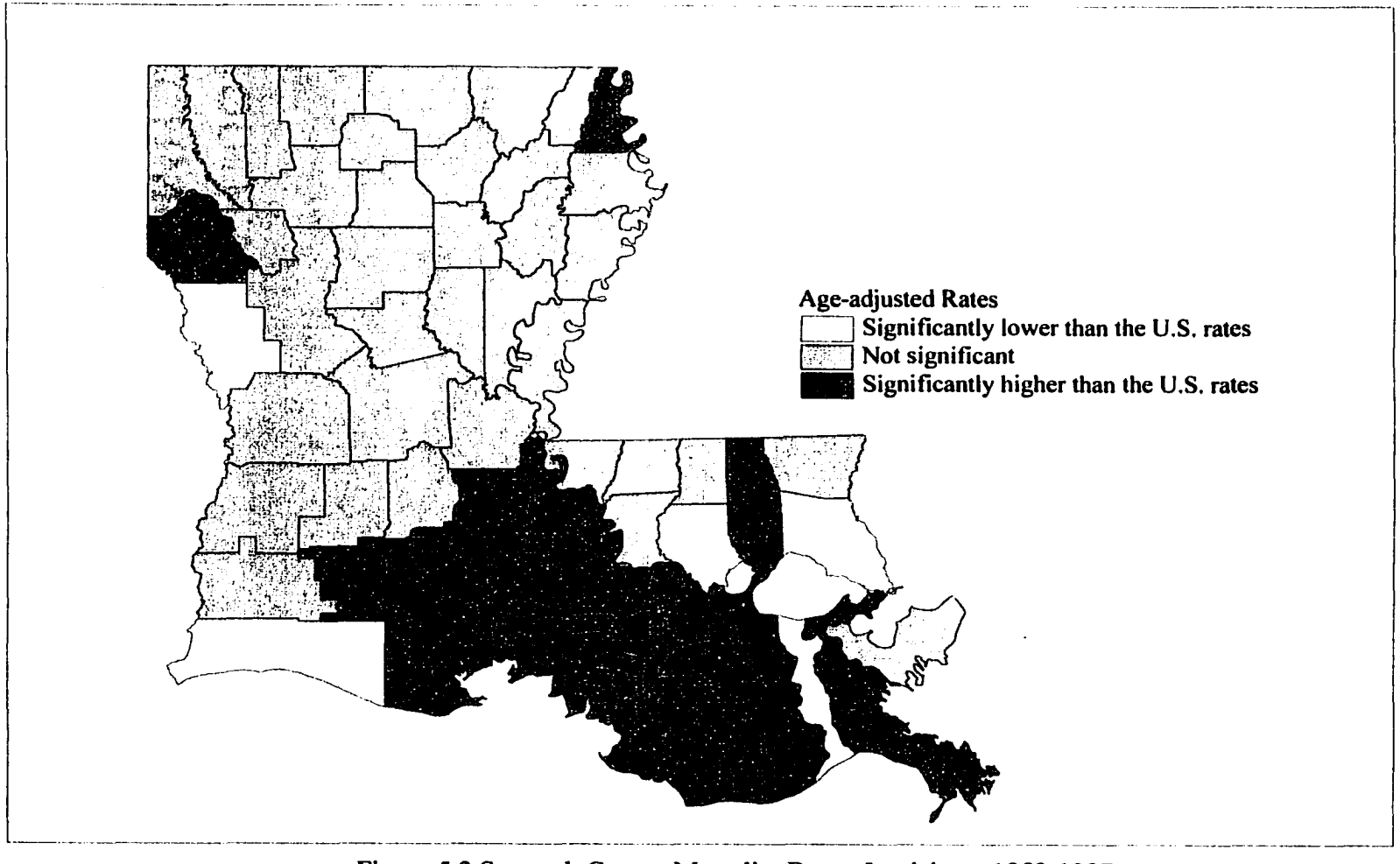

Figure 5.3 Stomach Cancer Mortality Rates, Louisiana: 1953-1987
(Source: Modified by Author from National Technical Information Service, 1992)

males than the U.S. (Table 5.3). Their distributions were concentrated west of the Mississippi River in southern Louisiana.

Colorectal Cancer: In case of colorectal cancer, cancer mortality rates for South Louisiana were either similar to, or lower than those for the U.S. rates (Appendices O-P). The only exceptions were that Orleans and St. Bernard for white males and Orleans for white females and nonwhites were significantly elevated when compared to the U.S. rates.

Breast Cancer: Even though breast cancer is the most common cancer in Louisiana women, cancer mortality rates for white females in South Louisiana were either similar to, or lower than those of the U.S. Orleans was the only parish in which breast cancer rates for nonwhites were significantly higher than those of the U.S. (Appendix Q).

Prostate Cancer: Prostate cancer is more prevalent in nonwhites than in whites (Appendix R). Mortality rates for prostate cancer in South Louisiana were similar or significantly lower than those of the U.S. Significantly higher rates were observed for Vermilion for white males, and for Orleans for nonwhite males.

In general, cancer mortality rates for most parishes in South Louisiana were either statistically significantly higher than, or similar to, the combined rates in the U.S. rates for cancers of all sites combined, lung, and stomach. The results of significance test for cancer mortality rates from 1953 to 1987 were relatively similar to previous studies and cancer incidence reports. Previous cancer mortality studies showed that the following cancer mortality rates were significantly higher than those of the U.S: all sites for males in Louisiana during 1950-1967; lung cancer for males in Louisiana during

1930-1932; lung cancer for white males in parishes of South Louisiana during 1950-1954, 1970-1975 (Correa et al. 1983), and 1970-1980 (Pickle et al. 1987); stomach for nonwhites during 1950-1967 (Correa et al. 1983; Pickle et al. 1990).

In the case of incidence studies in South Louisiana during 1983-1987, as compared to the SEER rates, lung cancer in white males (Chen et al. 1992a) and stomach cancer in black males (Fontham et al. 1992) showed significantly higher rates. In contrast to the high death rates in South Louisiana, incidence rates for all cancers combined were either the same as, or lower than, the SEER rates. The geographic distributions of the major cancer mortality rates and their relationships with environmental variables in Louisiana are further studied and discussed in the following sections.

5.2 Factor Analysis of Cancer Mortality Rates in Louisiana

Within Louisiana parishes, factor analysis was conducted on the 16 cancer mortality rates from 1953 to 1977, from 1978 to 1987, and from 1953 to 1987, in order to find a parsimonious explanation of the geographic patterns of the major cancer mortality rates, examine the change of the patterns of cancer mortality rates in two time periods (1953-1977 and 1979-1987), and compare with the geographic patterns of the U.S (section 4.3). This section explains the results of factor analysis of cancer mortality rates in Louisiana from 1953 to 1977 and 1978 to 1987.

5.2.1 From 1953 to 1977

Factor analysis was performed on 16 cancer mortality rates in Louisiana by parish from 1953 to 1977 after the sampling adequacy of the factor model was evaluated. Initial eigenvalues and a Scree plot suggested six common factors,

accounting for 68.1% of the variance. Ten factors were dropped because they did not meet the desirable criterion that an eigenvalue is equal to or greater than one loading on the factors. Final factor analysis was conducted on the remaining 6 factors and these factors were labeled.

Table 5.4 presents the results from factor analysis. Lung cancer mortality rates for nonwhite males and for nonwhite females generated a communality of 0.88 and 0.80, the highest of all. The lowest values were generated by lung cancer mortality rates for white females (0.20), indicating only a weak association of this cancer type with the factors.

The six factors produced finally explained about 50.6% of the total variance among the variables. Factor 1 has an eigenvalue of 3.354, which equals 21.0 % of the total variance. On account of the high degree of variance it explains, factor 1 was considered a pervasive influence on other factors in the analysis. Factor 2 explains 9.7 % and factor 3 does 7.0 % of the total variance. The remaining 3 factors together account for only 12.9 % of the variance.

The varimax rotation method was invoked to perform an orthogonal factor rotation on the six factors (Table 5.5). This procedure produced the simplest factor structure with factors containing significant loadings for a few variables and small loadings for the rest of the variables in the analysis. The factor loading indicates the degree of correlation between the factor and the individual variable. The highest loading variable in each factor was chosen to represent the information conveyed in that factor. Factor 1 represented a general factor with 4 of the 16 cancer variables loading at 0.5 or higher: nonwhite male stomach (0.76), white female colorectum (0.61), white

Table 5.4 Statistics for Factor Analysis of Cancer Mortality Rates in Louisiana for 1953-1977

Variable	Communality	*	Factor	Eigenvalue	Pct of Var	Cum Pct
BNWF5377	.4264	*	1	3.354	21.0	21.0
BWF5377	.7427	*	2	1.559	9.7	30.7
CRNWF5377	.3328	*	3	1.115	7.0	37.7
CRNWM5377	.2683	*	4	.862	5.4	43.1
CRWF5377	.5341	*	5	.655	4.1	47.2
CRWM5377	.6816	*	6	.548	3.4	50.6
LNWF5377	.7990	*				
LNWM5377	.8805	*				
LWF5377	.2016	*				
LWM5377	.4298	*				
PNWM5377	.5486	*				
PWM5377	.3243	*				
SNWF5377	.4339	*				
SNWM5377	.6770	*				
SWF5377	.4499	*				
SWM5377	.3628	*				

BNWF5377 - Nonwhite female breast cancer mortality from 1953 to 1977
BWF5377 - White female breast cancer mortality from 1953 to 1977
CRNWF5377 - Nonwhite female colorectal cancer mortality from 1953 to 1977
CRNWM5377 - Nonwhite male colorectal cancer mortality from 1953 to 1977
CRWF5377 - White female colorectal cancer mortality from 1953 to 1977
CRWM5377 - White male colorectal cancer mortality from 1953 to 1977
LNWF5377 - Nonwhite female lung cancer mortality from 1953 to 1977
LNWM5377 - Nonwhite male lung cancer mortality from 1953 to 1977
LWF5377 - White female lung cancer mortality from 1953 to 1977
LWM5377 - White male lung cancer mortality from 1953 to 1977
PNWM5377 - Nonwhite male prostate cancer mortality from 1953 to 1977
PWM5377 - White male prostate cancer mortality from 1953 to 1977
SNWF5377 - Nonwhite female stomach cancer mortality from 1953 to 1977
SNWM5377 - Nonwhite male stomach cancer mortality from 1953 to 1977
SWF5377 - White female stomach cancer mortality from 1953 to 1977
SWM5377 - White male stomach cancer mortality from 1953 to 1977

(Source: Calculated by author from NTIS cancer mortality rates, 1992)

Table 5.5 Rotated Factor Matrix in Louisiana for 1953-1977

	Factor 1	Factor 2	Factor 3
SNWM5377	.7598	.2038	.2203
CRWF5377	.6079	- .0860	.0216
CRWM5377	.5908	.1526	.0527
BNWF5377	.5216	- .1763	.1305
LNWM5377	.0859	.9172	.1518
CRNWM537	-.0325	.4806	- .1259
LWM5377	.3495	.3605	.1292
SWF5377	.0552	- .0291	.6443
SWM5377	.1704	.0230	.5325
SNWF5377	.3861	.0917	.4986
PNWM5377	- .1468	.0544	.4409
BWF5377	.0505	.3677	.1361
CRNWF537	.1887	- .0906	.2255
LNWF5377	.0745	.5144	.3794
LWF5377	.0139	.1768	.0515
PWM5377	.0140	- .0540	-.0692

See Table 5.4 for footnotes.
(Source: Calculated by author from NTIS cancer mortality rates, 1992)

male colorectum (0.59), and nonwhite female breast (0.52). Since stomach cancer in nonwhite males and colorectal cancer in whites had the highest variable loadings on factor 1, they were selected to represent the information that this factor conveys.

Factor scores of the first factor in Louisiana were demonstrated in Figure 5.4. Parishes which had high scores on this factor had higher than average mortality rates from nonwhite male stomach, white male and female colorectum, and nonwhite female breast cancers. Parishes which had low scores on this factor tended to have lower than average mortality rates from those cancers mentioned above. The highest were in St.

Figure 5.4 Factor Score Distribution for Factor 1, Louisiana (1953-1977)
(Source: Calculated by Author from National Technical Information Service, 1992)

Bernard, 3.7, and St. John the Baptist, 1.6, in southeastern Louisiana. Higher scores (greater than 1.0) emerged in St. James, Assumption, and Orleans in southeastern Louisiana and Jefferson Davis in the southwestern Louisiana. Relatively high and far more modest scores prevailed in some parishes (Iberia, St. Mary, and Plaquemines) of the lower Mississippi River, Caldwell in the northern Louisiana, and Avoyelles in central Louisiana. Also, the west side of Mississippi River showed all positive scores. Lower scores were located predominantly in most of northern and western Louisiana. Areas of the lowest scores developed in Cameron, -3.30, and West Feliciana, -1.08. Generally, major cancers representing factor 1 were related to cancers of digestive system, and this factor appeared to represent some regions of lower southeastern Louisiana.

Lung cancer for nonwhite males (0.92) and females (0.51) loaded highly onto factor 2. Others, except for cancers listed above, had a low variable loading on this factor. In particular, this factor was negatively correlated with at least one variable of cancer types (excluding lung cancer): breast cancer for nonwhite females, colorectal cancer for whites and nonwhite females, prostate cancer for white males, and stomach cancer for white females.

The highest scores on factor 2 were found in Cameron (4.72), and Jefferson (1.56) of southern Louisiana (Figure 5.5). In particular, the geographic distributions of most high scores were concentrated in southern Louisiana. The lowest scores were generated by Caldwell (-1.97), Bienville (-1.54), and Red River (-1.40) in northern Louisiana. Most northern parishes developed low and negative scores, except for Madison which had a positive score. Lung cancer mortality rates for nonwhites had the

Figure 5.5 Factor Score Distribution for Factor 2, Louisiana (1953-1977)
(Source: Calculated by Author from National Technical Information Service, 1992)

highest loadings on factor 2, and this factor represented lung cancer mortality rates for nonwhites in southern Louisiana, which covers most wetlands.

Stomach cancer for white females (0.64) and males (0.53) loaded highest onto factor 3. Generally, the third factor did not have higher coefficients on the variables, compared with the first and second factor. This factor appeared to represent stomach cancer mortality rates. The highest loading variable in each factor was chosen to represent the information conveyed in that factor. But following Cleverley and Nutt (1984), variables were considered to load on a factor if they had positive coefficients equal to or greater than 0.65.

The highest scores of factor 3 were in Lafourche (1.65) and St. Landry (1.49) and the lowest scores were in Grant (-1.99) and Concordia (-1.50) (Figure 5.6). Relatively higher and moderate scores prevailed west of the Mississippi River in southern Louisiana, except for Assumption and Jefferson. The areas extended from west of the Mississippi River in southern Louisiana through northeastern parishes. Areas of lower scores also developed in northern Louisiana as those of factor 1 did. Thus, a major cancer representing factor 3 was stomach, and this factor was widely distributed west of the Mississippi River.

Factor 4 was best represented by breast cancer for white females (factor loading = 0.74). Factor 5 and factor 6 were the only factor onto lung cancer for white females (factor loading = 0.59) and prostate cancer for white males (factor loading = 0.56), respectively. But 4, 5, and 6 factors with an eigenvalue less than one were not explained for this study.

Figure 5.6 Factor Score Distribution for Factor 3, Louisiana (1953-1977)
(Source: Calculated by Author from National Technical Information Service, 1992)

Also, other factor analysis methods (principal components, maximum likelihood, and unweighted least squares) were performed and these results were then compared. The procedures produced similar factor groups of variables.

5.2.2 From 1978 to 1987

The same procedure of factor analysis for the 16 cancer mortality rates in Louisiana from 1978 to 1987 was undertaken to compare with that for 1953 to 1977. The 0.62 value of the KMO measure for sampling adequacy of factor model allowed a factor analysis to proceed.

Final statistics of factor analysis (Table 5.6) show communalities, the proportions of variance accounted for by common factors. The cancers of white male colorectum (0.77), white female colorectum (0.74), nonwhite female stomach (0.69), and nonwhite female colorectum (0.67) developed the highest communality among the 16 types. White female breast (0.19) and white female stomach (0.27) showed the lowest communality value.

Table 5.6 shows that the final six factors extracted explain 50.7% of the total variance. Factor 1 has a variance of 3.14, which amounts to 19.7% of the total variance. Since it explains one fifth of the total variance, this factor provides the most important pattern of geographical distribution. Factors 2, 3, and 4 explain 8.5%, 8.2%, and 6.5% of the variance, respectively. However, the eigenvalues of factor 5 and 6 are too small (less than one), and the factors might not be meaningful. Therefore, this study only explained 1, 2, 3, and 4 factors.

Rotated factor matrix (Table 5.7) indicates that the first factor showed a strong positive correlation with cancers of nonwhite male stomach (0.69), nonwhite female

Table 5.6 Statistics for Factor Analysis of Cancer Mortality Rates in Louisiana for 1978-1987

Variable	Communality	*	Factor	Eigenvalue	Pct of Var	Cum Pct
BNWF7887	.5947	*	1	3.148	19.7	19.7
BWF7887	.1907	*	2	1.361	8.5	28.2
CRNWF7887	.6765	*	3	1.312	8.2	36.4
CRNWM7887	.5772	*	4	1.039	6.5	42.9
CRWF7887	.7472	*	5	0.741	4.6	47.5
CRWM7887	.7682	*	6	0.513	3.2	50.7
LNWF7887	.6079	*				
LNWM7887	.6359	*				
LWF7887	.3740	*				
LWM7887	.3228	*				
PNWM7887	.4242	*				
PWM7887	.3998	*				
SNWF7887	.6977	*				
SNWM7887	.5133	*				
SWF7887	.2757	*				
SWM7887	.3082	*				

BNWF7887 - Nonwhite female breast cancer mortality from 1978 to 1987
BWF7887 - White female breast cancer mortality from 1978 to 1987
CRNWF7887 - Nonwhite female colorectal cancer mortality from 1978 to 1987
CRNWM7887 - Nonwhite male colorectal cancer mortality from 1978 to 1987
CRWM7887 - White male colorectal cancer mortality from 1978 to 1987
LNWF7887 - Nonwhite female lung cancer mortality from 1978 to 1987
LNWM7887 - Nonwhite male lung cancer mortality from 1978 to 1987
LWF7887 - White female lung cancer mortality from 1978 to 1987
LWM7887 - White male lung cancer mortality from 1978 to 1987
PNWM7887 - Nonwhite male prostate cancer mortality from 1978 to 1987
PWM7887 - White male prostate cancer mortality from 1978 to 1987
SNWF7887 - Nonwhite female stomach cancer mortality from 1978 to 1987
SNWM7887 - Nonwhite male stomach cancer mortality from 1978 to 1987
SWF7887 - White female stomach cancer mortality from 1978 to 1987
SWM7887 - White male stomach cancer mortality from 1978 to 1987

(Source: Calculated by author from NTIS cancer mortality rates, 1992)

Table 5.7 Rotated Factor Matrix in Louisiana for 1978-1987

	Factor 1	Factor 2	Factor 3	Factor 4
SNWM7887	.6889	.0409	.0906	-.0322
CRNWF788	.6835	.0143	.2885	-.2196
BNWF7887	.6518	.2345	.0275	.2802
LWM7887	.3736	.0864	.3325	.0545
LNWF7887	.0623	.7168	.0373	.0747
SNWF7887	-.0699	.6962	.2494	-.2802
LNWM7887	.2975	.6271	.0727	.3792
SWM7887	.3513	.3911	-.1539	-.0823
CRWF7887	.0096	.0402	.8480	.1444
LWF7887	.3301	.1044	.4872	-.0933
SWF7887	.1922	-.1290	.2488	.2369
PNWM7887	-.0657	-.0200	.0209	.6460
CRWM7887	.0512	.2565	.1102	.1156
PWM7887	-.2664	.0085	.2231	.0840
CRNWM788	.0301	.1827	.3005	.4479
BWF7887	.1796	.0293	-.0400	0167

See Table 5.6 for footnotes.
(Source: Calculated by author from NTIS cancer mortality rates, 1992)

colorectum (0.68), and nonwhite female breast cancer (0.65). The second factor was highly correlated with cancers of nonwhite female lung (0.72), nonwhite female stomach (0.70), and nonwhite male lung cancer (0.63). But the first factor showed a slightly negative correlation with prostate cancer, and the second factor also showed a slightly negative or no correlation with this cancer. The third factor and the fourth factor were highly correlated with only white female colorectal cancer (0.85) and nonwhite male prostate cancer (0.65), respectively. But these factors did not show a

high correlation with other cancers (except for white female colorectal cancer and nonwhite male prostate cancer respectively) and in particular, the factor 4 had a weakly negative or little correlation with stomach cancers.

Factor scores of the four factors were demonstrated in four maps of Louisiana shown in Figures 5.7-5.10. A wide range of scores developed on the first factor (Figure 5.7). The highest scores of factor 1 were in Cameron (3.90), Assumption (1.70), and St. Bernard (1.40). High scores were more dispersed. Relatively high scores prevailed in the lower southern parishes, except for Vermilion (-0.53), Terrebonne (-0.39), and Jefferson (-0.06). Far more modest scores and low scores showed in most of the northeastern and middle parishes. Areas of the lowest scores developed in East Feliciana (-1.82), Catahoula (-1.57), and West Carroll (-1.57) in the northeastern Louisiana as those of low scores generated in northeastern Louisiana.

In the second factor (Figure 5.8), St. Bernard (2.10), Livingston (1.97), Terrebonne (1.83), and St. Mary (1.63) had the highest scores and most of southern Louisiana showed relatively high scores, except for Cameron. Most parishes in northern Louisiana developed moderate to low scores, with the exception of Caldwell (0.61). The lowest scores were in Cameron (-2.83), Bienville (-1.69), and Catahoula (-1.53).

A relatively narrow range of scores developed on the third factor (Figure 5.9). The highest scores on factor 3 were in East Carroll (2.91), Catahoula (1.92), St. James (1.67), and St. Bernard (1.53). The lowest scores were in St. Helena (-1.83), West Baton Rouge (-1.70), and La Salle (-1.58). High scores were dispersed in parishes near to the state boundary. Most of central Louisiana developed moderate to low scores.

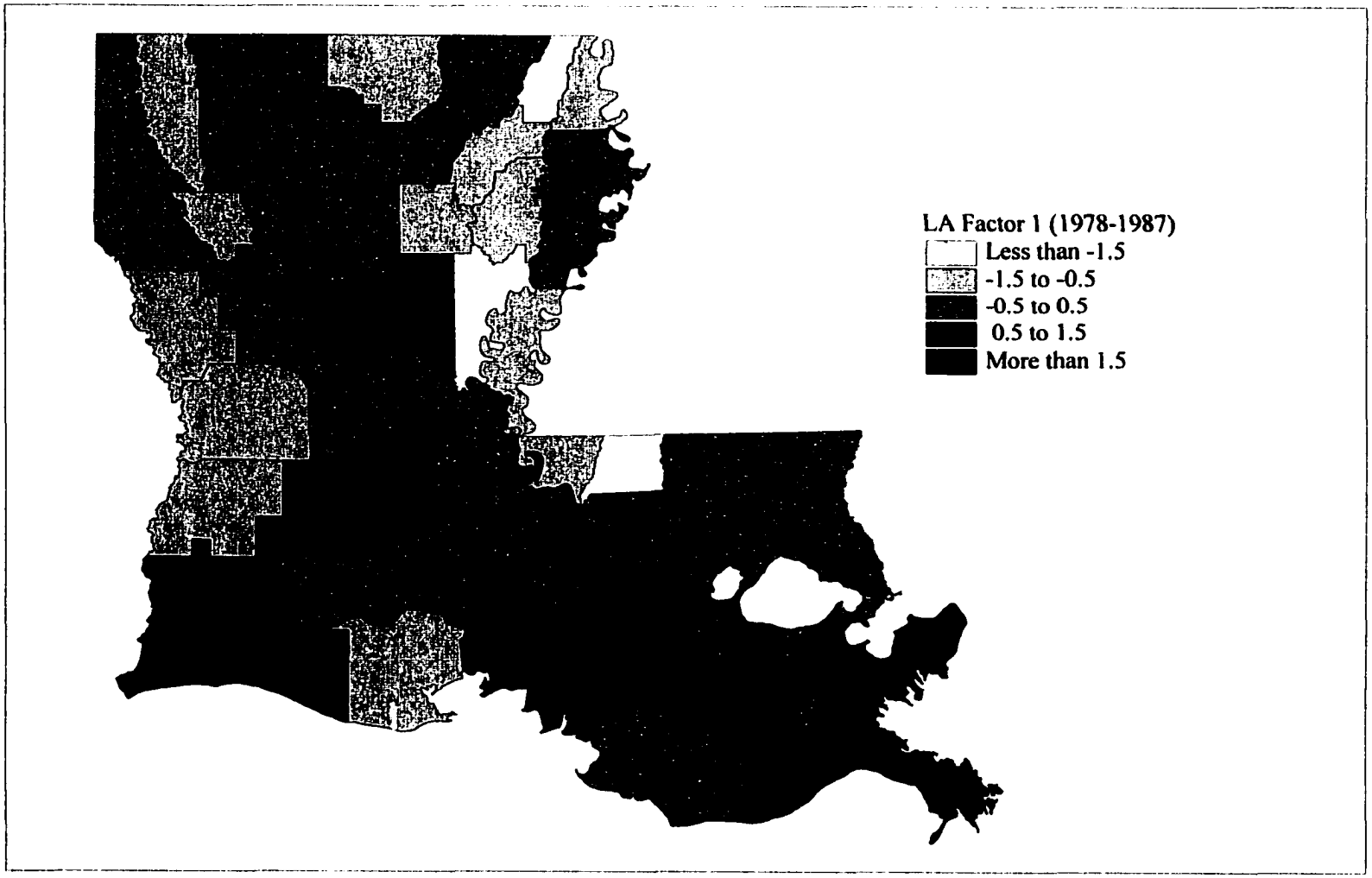

Figure 5.7 Factor Score Distribution for Factor 1, Louisiana (1978-1987)
(Source: Calculated by Author from National Technical Information Service, 1992)

Figure 5.8 Factor Score Distribution for Factor 2, Louisiana (1978-1987)
(Source: Calculated by Author from National Technical Information Service, 1992)

Figure 5.9 Factor Score Distribution for Factor 3, Louisiana (1978-1987)
(Source: Calculated by Author from National Technical Information Service, 1992)

The highest score of factor 4 was in La Salle (2.39) and the lowest scores were Cameron (-2.62), St. Bernard (-2.10), and Red River (-1.61) (Figure 5.10). Parishes which had high or low scores on this factor had higher or lower rates than average mortality rates from nonwhite male prostate cancer. High scores were presented in some northern (East Carroll, Caddo, De Soto, and Madison) and southern parishes (St. James, Jefferson Davis, and, Assumption). Most parishes developed moderate to low scores. The geographic distributions of these scores were more fragmented than those of nonwhite male prostate cancer of higher rates. In fact, the geographic distributions of factor scores of factor 3 and 4 were very similar to those of mortality rates of colorectal cancer for white females and prostate cancer for nonwhite males, respectively.

Finally, interpretation of the factors appeared generally possible. The first factor could not be easily interpreted but the regions of high factor score were usually distributed in lower southern Louisiana. Major cancers representing factor 1 were related to cancers of the digestive system since stomach cancer for nonwhite males and colorectal cancer for whites had the highest variable loading on factor 1. In the second factor, lung cancer mortality rates for nonwhites had the highest loadings and the factor was highly concentrated in lower southeastern Louisiana. Therefore, the second one could be considered to be lung cancer for nonwhites in the southern parishes. The major cancer representing the third factor was colorectum in white females, and high scores for this factor were dispersed in parishes near the state boundary. This fourth factor appeared to represent mortality rates from prostate cancer for nonwhite males, and positive factor scores for this factor were relatively concentrated in the central Louisiana.

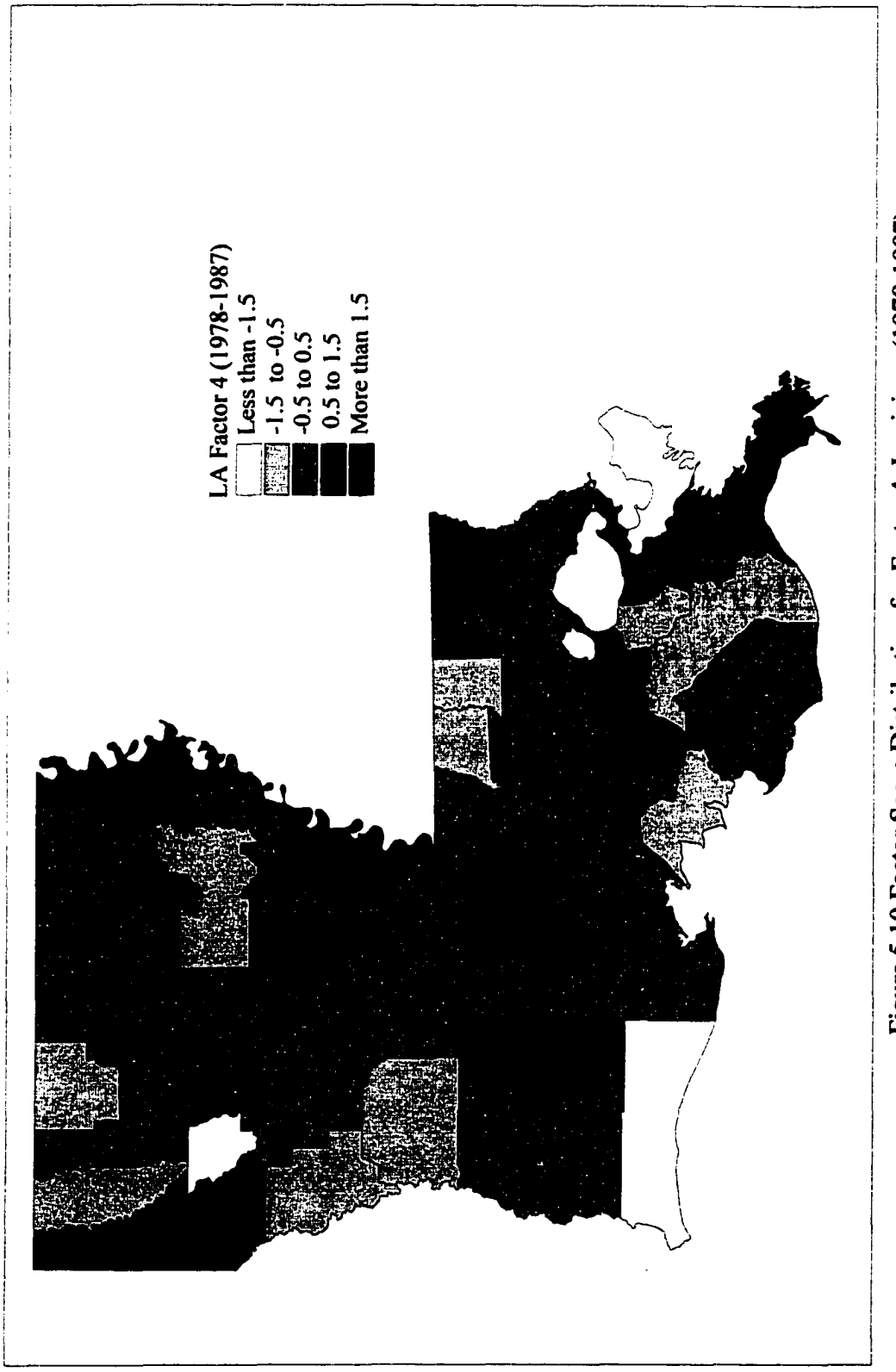

Figure 5.10 Factor Score Distribution for Factor 4, Louisiana (1978-1987)
(Source: Calculated by Author from National Technical Information Service, 1992)

Even though factor 5 and 6 (with eigenvalue less than one) were not explained here, factor 5 was best represented by the colorectal cancer for white males (factor loading =0.75) and factor 6 was the only factor onto colorectal cancer for nonwhite males (factor loading = 0.56).

5.2.3 Factor Analysis of Cancer Mortality Rates in Louisiana

Table 5.8 summarizes major cancer sites related to extracted factors in the U.S. and Louisiana. As mentioned in Chapter 4.3, major factors in the U.S for 1953-1987 could be interpreted as "lung cancers for whites" (for factor 1) and "cancers of colorectum and breast" for whites (for factor 2) and nonwhites (for factor 3). Within Louisiana, major factors extracted from factor analysis during the two time periods were relatively similar. Major cancer sites related to extracted factors were cancers of the digestive system (for factor 1) and lung cancers for nonwhites (for factor 2).

In Louisiana, stomach and colorectal cancers representing factor 1 are the most common cancers of the digestive system. Stomach cancer mortality rates are more than twofold higher in nonwhites than whites. Stomach cancer mortality rates among nonwhites in parishes of South Louisiana were significantly higher than those of the U.S. On the other hand, colorectal cancer rarely occurred in South Louisiana as well as Louisiana, as compared to other areas of the U.S. Only Orleans had the highest age-adjusted rates for colorectal cancer, and was significantly higher than the U.S.

Cancers of gastorintestinal tract, representing the first factor in Louisiana (unlike the first factor of the U.S), might be partly explained by the dietary habits for certain cultures and socioeconomic factors. Studies of gastric cancer etiology in South Louisiana suggested a major role for dietary determinants (Correa et al. 1985b; Fontham

Table 5.8 Major Cancer Sites Related to Extracted Factors in Louisiana and the U.S.

Factor	U.S.: 1953-1987	LA:1953-1977	LA:1978-1987
Factor 1 (>0.5)	White male lung (0.94) White female lung (0.56)	Nonwhite male stomach (0.76) White female colorectum (0.61) White male colorectum (0.59) Nonwhite female breast (0.52)	Nonwhite male stomach (0.69) Nonwhite female colorectum (0.68) Nonwhite female breast (0.65)
Factor 2 (>0.5)	White male colorectum (0.78) White female breast (0.56) White female colorectal (0.52)	Nonwhite male lung (0.92) Nonwhite female lung (0.51)	Nonwhite female lung (0.72) Nonwhite female stomach (0.70) Nonwhite male lung (0.63)
Factor 3 (>0.5)	Nonwhite female breast (0.74) Nonwhite female colorectum(0.64)	White female stomach (0.64) White male stomach (0.53) Nonwhite female stomach (0.50)	White female colorectum (0.85)
Factor 4 (>0.5)			Nonwhite male prostate (0.65)

(Source: Calculated by author from NTIS cancer mortality rates, 1992)

et al. 1992). Diets low in fresh fruits and vegetables and high in gastric irritants, such as salt, were associated with high gastric cancer risk. Socioeconomic status was known to influence gastric cancer risk (Normura 1982). However, international studies showed that countries with high rates of stomach cancer, such as Japan, China, and Costa Rica, tended to have low rates of colorectal cancers (Muir et al. 1987).

Cigarette smoking is considered the major cause of lung cancer (Chen et al. 1992a). Lung cancer rates are three times higher in males than in women. White males in Louisiana have shown significantly higher rates than those of the U.S. However, the major cancers representing Louisiana factor 2 were lung cancer for nonwhites. Besides tobacco smoking, the high density of industries and socioeconomic factors related to nonwhite population in Louisiana might partly explain this factor.

The basic difference of extracted factors between the two time periods lay in the change of high loading of stomach cancer (of factor 1 and 3 in 1953-1977) and of lung cancer (from nonwhite males to nonwhite females for factor 2). It indicated the importance of stomach cancer mortality rates in the earlier period, but it has since been declining steadily.

The factor scores maps show that high scores of factor 1 have been extended from southeastern Louisiana to southern Louisiana, whereas those of factor 2 have been shifted from southern Louisiana to southeastern Louisiana. In general, the geographical distributions of major factor scores were more prominent in southern Louisiana along the west bank of the Mississippi River. They also have increased in the northern part of Louisiana and widely extended to western parishes of the state. The spatial patterns of cancer mortality might change over time partly because of changes in environmental

factors, such as population density, racial composition, per capita income, environmental hazards, cigarette smoking, alcohol consumption, culture and dietary habits, the use of chlorinated water from the Mississippi River, occupation, socioeconomic status, stress, and medical practices. Therefore, the stepwise regression method in section 5.5 are employed to provide information concerning possible etiology of disease.

The application of factor analysis for this study could be useful in providing information on the hypothetical factors that may be closely connected with the onset of various diseases. The results obtained by factor analysis indicated that cancer sites form some clusters with respect to their geographical distributions. These clusters, if detected with high significance, might be used to postulate the existence of some common causes of cancer surrounding the clusters.

In summary, the first factor represented cancers related to the digestive system (such as nonwhite male stomach and female colorectum) during the two periods, but the factor regions developed in southeastern Louisiana for 1953-1977 and in southern Louisiana for 1978-1987. The second factor for both periods was considered to be the lung cancer mortality rates for nonwhites, and the factor regions were defined as in southern Louisiana for 1953-1977, and as in southeastern Louisiana for 1978-1987. Major cancer sites of factor 3 in the earlier period were stomach cancers but these cancers did not have very high coefficients on the factor, compared with those of the first and second factor. These factor regions were widely distributed west of the Mississippi River. Factor 3, in the latter period, appeared to represent the colorectal cancer for white females, and high factor scores for this factor were dispersed in

parishes near the state boundary. This fourth factor (for 1978-1987) represented mortality rates of prostate cancer for nonwhite males, and moderately high factor scores were generally concentrated in central Louisiana.

In spite of different study periods, some clustering existed with respect to the geographical distributions of various sites of cancer mortality (i.e., lung cancer for whites and cancers related to the digestive system). Several parishes in South Louisiana or west regions of the lower Mississippi River had exceptionally high factor scores on important factors. That is, the second factor to represent lung cancer mortality rates for nonwhites featured an excessive cluster in southern Louisiana during the two periods. Stomach cancer mortality rates were more concentrated and distributed west of the Mississippi River. The results by factor analysis of the geographical distributions showed that cancer sites form some clusters with respect to their geographical distributions, pointing further the need for a cluster detection analysis.

5.3 Spatial Autocorrelation Analysis of Cancer Mortality Rates by Parish

The correlogram analysis technique was utilized to examine the spatial-temporal patterns of cancer mortality rates in Louisiana in 1953-1987, 1953-1977, and 1978-1987. The purpose is to determine if significant autocorrelation exists among the parishes in Louisiana in terms of their site, sex, race-specific cancer mortality rates for 1953-1987, 1953-1977, and 1978-1987. Using a two-tailed test with a significant level of 0.05 (0.01), a z value outside the range of 1.96 (2.58) was considered significantly spatially autocorrelated in either positive or negative direction. The results of the test for spatial autocorrelation according to both assumptions (normalization and randomization) presented negligible difference. This was due to the nearly normal

nature of the cancer mortality distribution. The results based on the normalization assumption are explained in this study.

Spatial correlograms provided good representations of the effect of spatial scale on spatial clustering or spatial autocorrelation. As mentioned before, correlograms were used because they are less affected than variogram by deviations from the normal distribution (Goodchild 1986). When correlogram curves are irregular and undulating, they indicate the existence of mixed spatial scale effects. Any departures from smooth decline in the correlogram would suggest the existence of other factors. If the curve declines and goes up after several lags (V shape curve), it implies that similarity of cancer mortality rates exists for parishes that are several parishes apart. Such similarity most likely occurs in parishes with big cities, which are often several parishes apart, and that the overall spatial patterns are heterogeneous.

The results of the first test for spatial autocorrelation among cancer sites are shown in Figures 5.11-5.13 and Table 5.9. The results demonstrated quite a wide variation of spatial autocorrelation among the cancer types (Figure 5.11). For 35 years from 1953 to 1987, cancers of all sites combined, lung, and stomach cancers exhibited the strongest degree of positive spatial autocorrelation, indicating a great clustering in groups comprised of contiguous parishes. Cancers such as breast and colorectum exhibited significant spatial autocorrelation at lag 1 or 2. Prostate cancer did exhibit no positive autocorrelation but negative autocorrelation at lag 8.

Table 5.9 shows the standardized 'I' values among the three time periods (1953-1977, 1978-1987, and 1953-1987). Through the times, I values of autocorrelation were relatively stable and not changed. Examining the results of the earlier period in

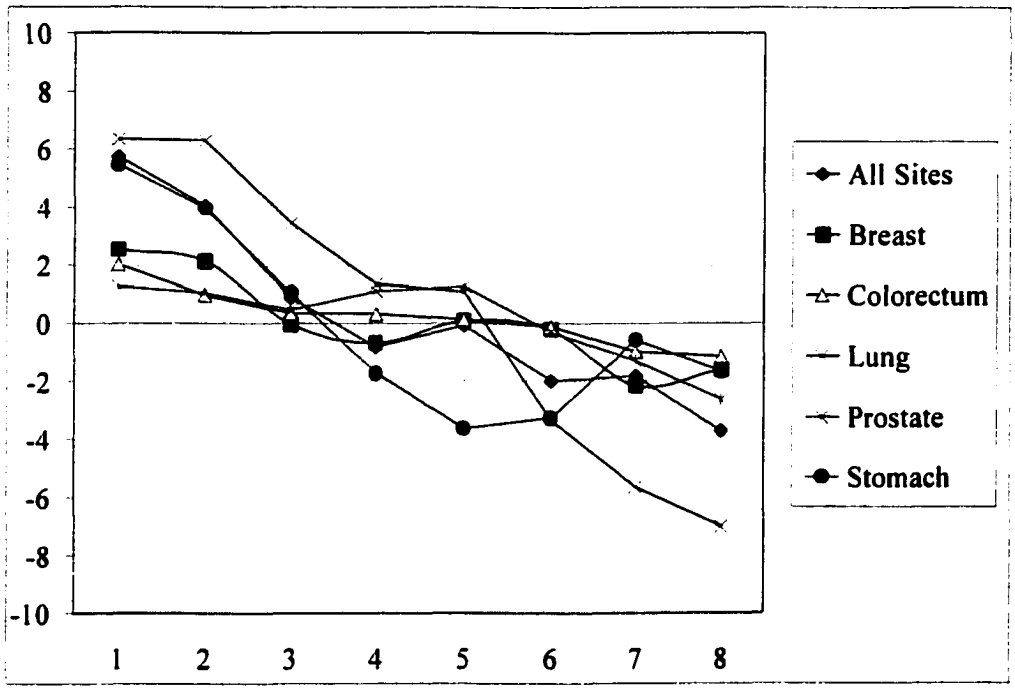

Figure 5.11 Correlograms for Selected Cancers, Louisiana, 1953-1987. For Figures 5.11-5.29, x axis shows spatial lags (1-8); y axis represents the z values at each lag; values outside the two dotted lines are considered statistically significant at $\alpha = 0.05$ level.

Table 5.9 Spatial Autocorrelation (Standardized I Values) Among Neighbors

Cancer Site	Period	1	2	3	4	5	6	7	8
All Sites	1953-87	5.7287**	4.0383**	0.8922	-0.8221	-0.0505	-1.9547	-1.8186	-3.6967**
	1953-77	5.2775**	3.2752**	-0.2613	-1.4965	-0.1673	-0.8582	-0.5583	-2.6341**
	1978-87	5.3498**	4.4149**	2.2713*	-0.1072	-0.7160	-3.2187**	-2.525*	-4.1758**
Breast	1953-87	2.5099*	2.1336*	-0.0142	-0.6787	0.0809	-0.2247	-2.1669*	-1.5877
	1953-77	2.6128**	2.1250*	-1.0794	-0.7071	0.4129	0.4426	-1.9279	-1.7655
	1978-87	2.2030**	1.1394	0.2359	-0.8702	-0.1362	-0.3792	-1.0535	-0.6598
Colorectum	1953-87	2.0243**	0.9795	0.3444	0.3136	0.1356	-0.110	-0.9589	-1.1245
	1953-77	1.6754	-0.1607	0.5266	0.2594	-0.9705	0.7743	-0.1063	-0.165
	1978-87	0.8303	2.9322**	0.3804	-0.7153	1.0535	-1.4861	-1.2345	-0.8935
Lung	1953-87	6.3706**	6.3057**	3.4574**	1.3406	1.0868	-3.3134**	-5.637**	-7.0132**
	1953-77	6.6024**	5.8206**	2.3126*	0.6383	1.6352	-2.0727*	-4.4084**	-7.0305**
	1978-87	5.1319**	4.9930**	2.8437**	0.5572	-0.9120	-3.5512**	-4.2635**	-3.6995**
Prostate	1953-87	1.2689	1.0578	0.4664	1.1141	1.2793	-0.2928	-1.2861	-2.6207**
	1953-77	0.7306	1.3580	0.2040	0.7888	1.2565	-1.2009	-0.054	-1.0649
	1978-87	1.1807	-0.1526	0.2053	0.2170	0.6448	1.2763	-1.1038	-2.7594**
Stomach	1953-87	5.4794**	3.9661**	1.0508	-1.7095	-3.6207**	-3.2439**	-0.5769	-1.6384
	1953-77	4.5116**	3.0921**	0.5483	-1.5698	-3.3288**	-2.1942*	-0.1001	-1.1864
	1978-87	5.1913**	5.9592**	2.2998*	0.3898	-2.0868*	-4.7993**	-4.3398**	-2.7536**

Standardized I Value by Normalization method; numbers 1-8 represent the neighbors from first-order to eight-order

*: Significant at 0.05 (two-tailed test)

** : Significant at 0.01 (two-tailed test)

comparison to those of the latter period explained the gain in autocorrelation by cancers of all combined sites and stomach (at lags 1, 2, and 3), and cancer of prostate (at lag 1), and the loss in autocorrelation by cancers of breast, lung, and colorectum (at lags 1 and 2).

To gain close insight into the influence of spatial scale on the cancer mortality patterns, spatial correlograms at the site, race, and sex-specific cancers were studied in detail.

5.3.1 Lung Cancer

For 1953-1987, lung cancer exhibited the strongest degree of positive and negative spatial autocorrelation. Its first three lags had significantly positive autocorrelation and its last three lags had negative autocorrelation. The shape of lung cancer correlogram, seen in Figure 5.11 is a strikingly declining line with increasing distance, except for a nearly horizontal line at the start of the graph extended from lag 1 to lag 2. In other words, approximately equal, but significantly high spatial autocorrelation existed among first-order, second-order, and third-order neighbors. The flatness of the first curve of the correlogram also indicated evidence of the existence of very large, multi-parish regions with similar lung cancer rates. Indeed, a typical lung cancer "cluster" has consisted of a third of the entire state of Louisiana since 1950s.

In two different time periods (1953-1977 and 1978-1987), the correlograms of lung cancer of both sexes among whites and nonwhites had a very similar declining curve from lag 1 to lag 4 and then 6 to 7, except for lags 5 and 8 (Figures 5.12 and 5.13). It meant similarity of the correlograms, indicating the spatial patterns of the earlier and latter periods are approximately alike.

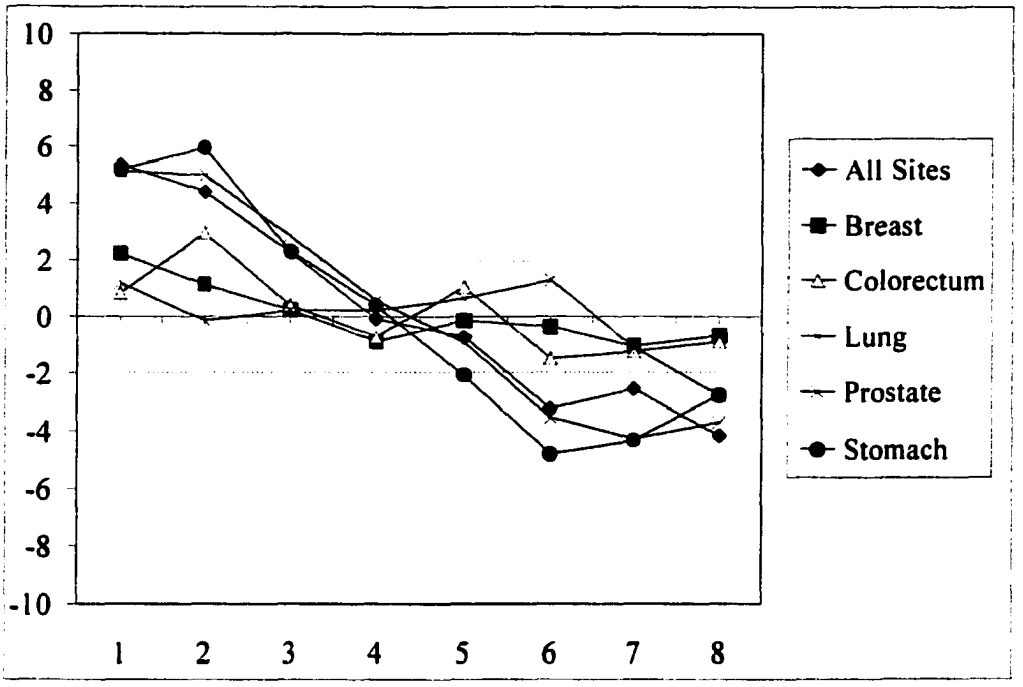

Figure 5.12 Correlograms for Selected Cancers, Louisiana, 1953-1977

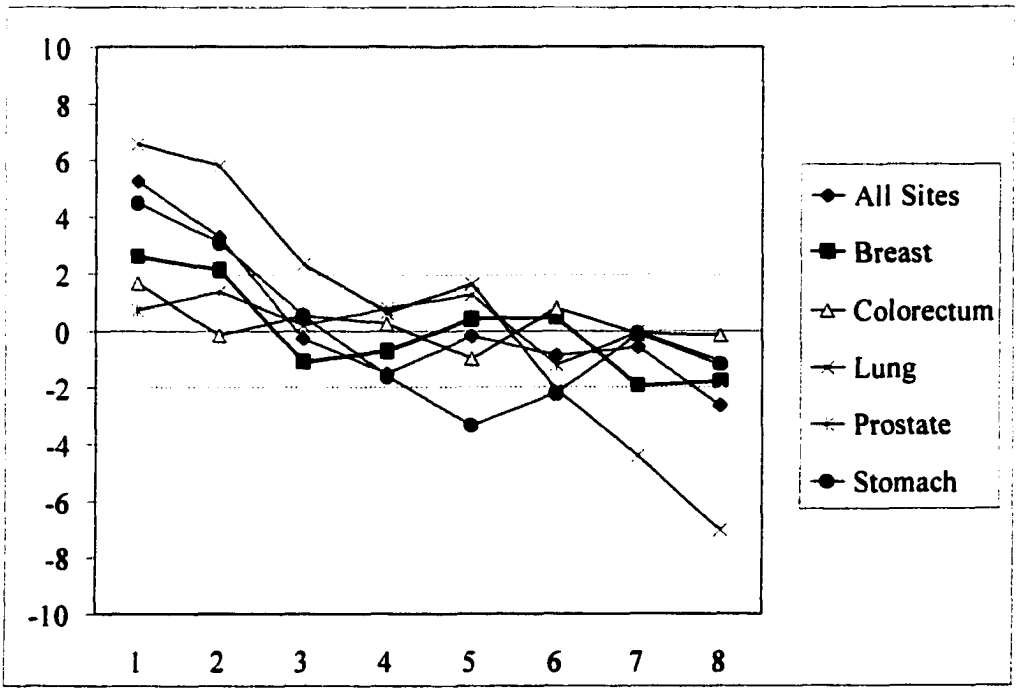

Figure 5.13 Correlograms for Selected Cancers, Louisiana, 1978-1987

As depicted in Figures 5.14 and 5.15, lung cancer in white males for 1953-1987 had a standard normal deviate of autocorrelation from -2.24 to 4.05 and the value in white female lung cancer ranged from –1.99 to 2.41. White male lung cancer had positive spatial autocorrelation at lags 1 and 2 and negative autocorrelation at lags 7 and 8, whereas white female lung cancer showed positive spatial autocorrelation at lag 1 and negative autocorrelation at lag 8. Generally, the correlograms of cancer mortality rates for whites showed an irregular V-shape.

Comparing the two periods, the spatial autocorrelation indices of cancer mortality rates for whites (both sexes) in 1953-1977 were high at lag 1, whereas those in 1978-1987 were below the statistically significant level and clearly exhibited no neighborhood effects. The form of correlograms for white male (white female) lung cancer had shifted from a big V-shape to an irregular or smooth curve (small V-shape) through the two periods. This indicated that the spatial patterns of high lung cancer mortality rates for whites were diversely extended from parishes along the west bank of the Mississippi River to parishes in northern Louisiana and included Cameron parish.

For 1953-1987, the standard normal deviate value in nonwhite male lung cancer ranged from -8.85 to 8.37, whereas in nonwhite female lung cancer it covered from -7.36 to 6.16 (Figures 5.16 and 5.17). Lung cancer among nonwhite males had the highest positive and negative spatial autocorrelation among the 16 cancer sites and the first three lags showed the highest z scores in the three periods. The form of the correlograms was most distinctive and consistent through time and the curves generally maintained the same form of decline with distance, except at lag 1. During the two time periods, a similarity in the correlograms was generally found. The latter period had

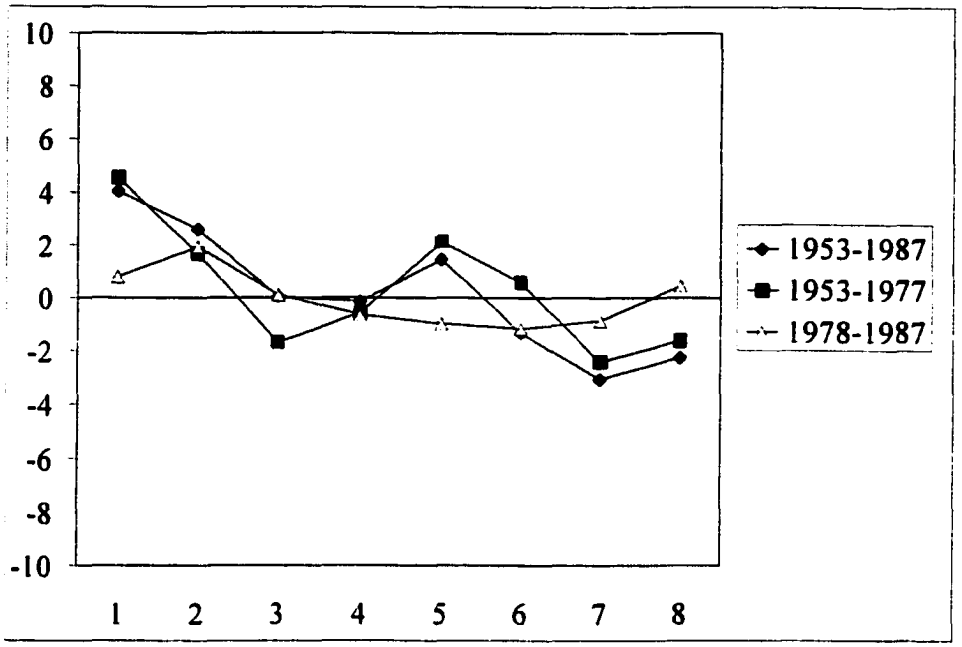

Figure 5.14 Correlograms for Lung Cancer (White Males)
Three different time periods: 1953-1987, 1953-1977, and 1978-1987

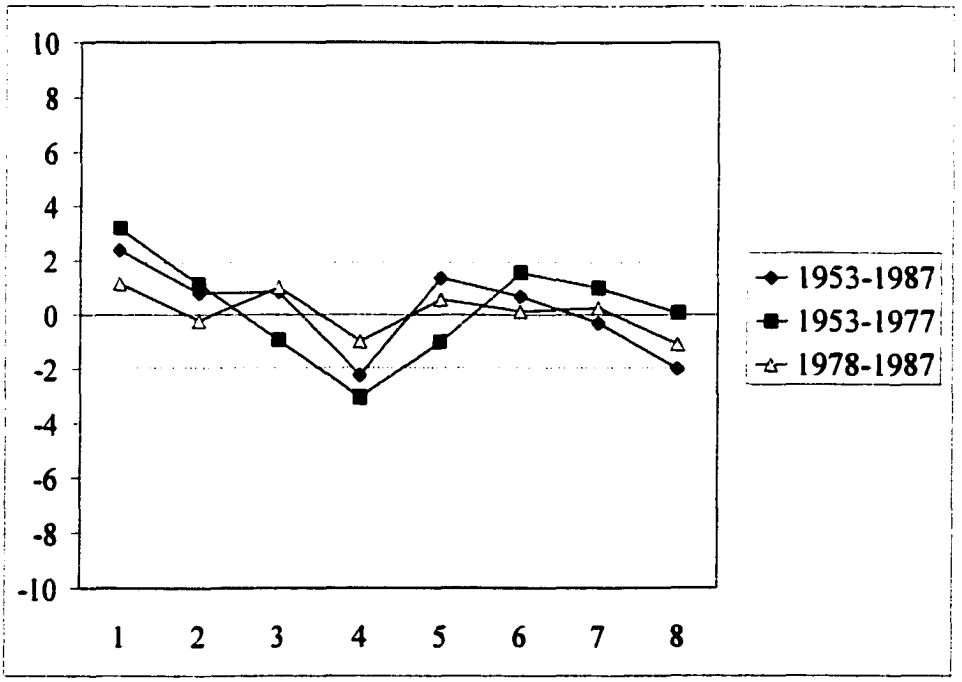

Figure 5.15 Correlograms for Lung Cancer (White Females)

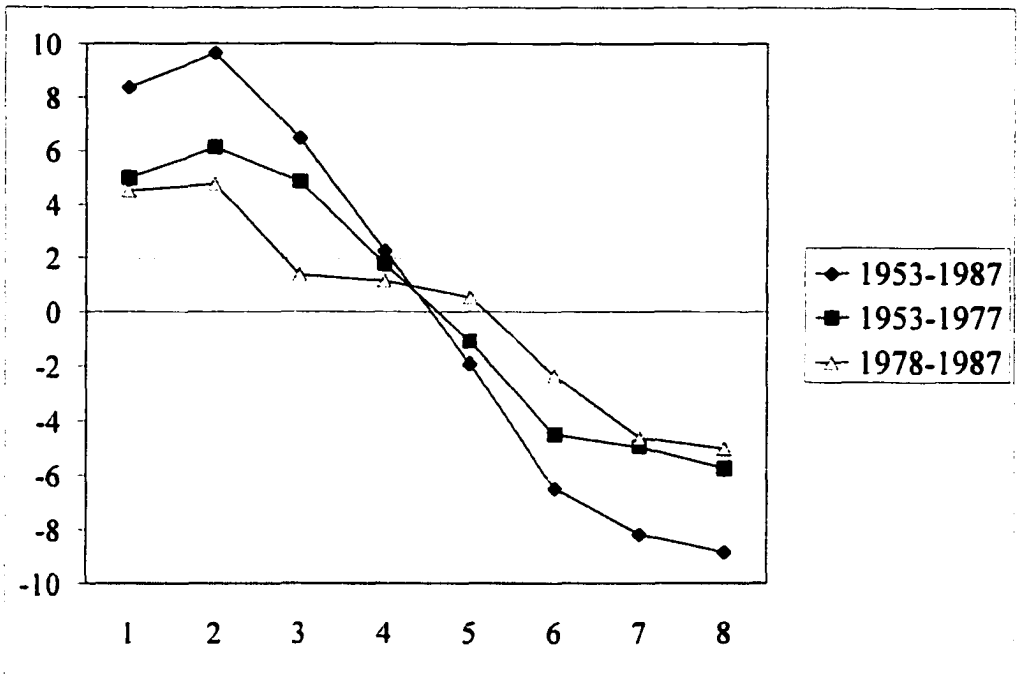

Figure 5.16 Correlograms for Lung Cancer (Nonwhite Males)

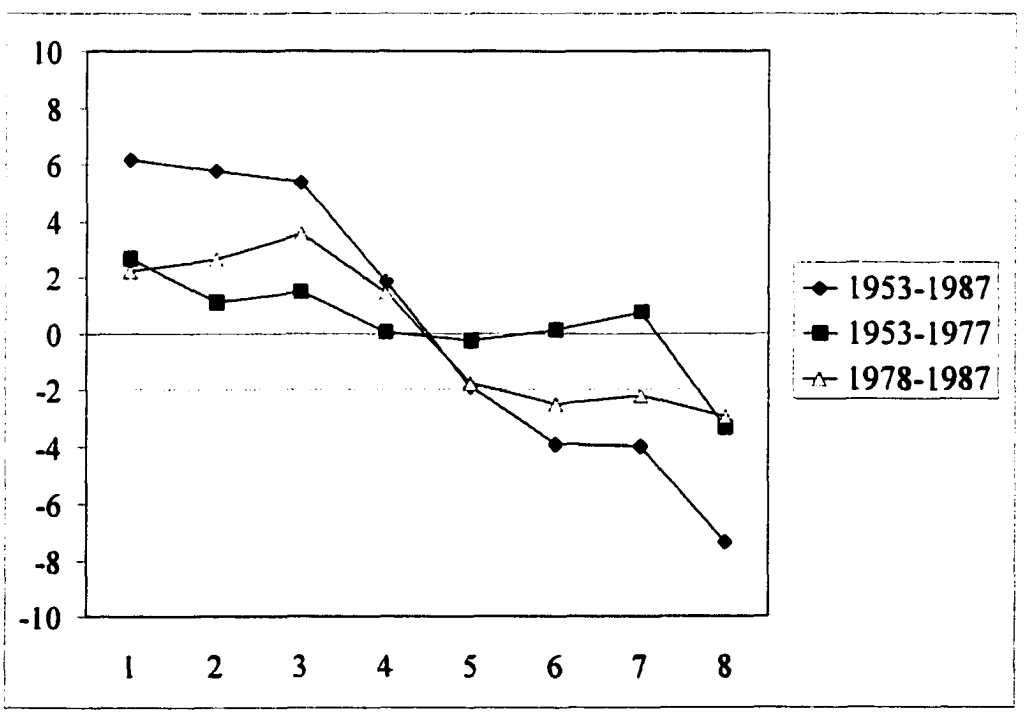

Figure 5.17 Correlograms for Lung Cancer (Nonwhite Females)

lower spatial autocorrelation indices than the earlier and in particular, their value in 1978-1987 decreased rapidly at spatial lag 3. This implied that the spatial patterns of high lung cancer mortality rates for nonwhite males have been changed but were still concentrated in southern Louisiana. The high spatial autocorrelation indicated a symmetrical pattern, with large patches of parishes having high cancer mortality rates serving as centers of spread to the neighboring parishes (Lam et al. 1996).

Lung cancer mortality rates among nonwhite females had the second highest positive and negative spatial autocorrelation among the 16 cancer sites. As with nonwhite male lung cancer, nonwhite females had the highest positive at lags 1, 2, and 3 and negative spatial autocorrelation at lags 6, 7, and 8. The curve of the correlograms decreased consistently through time. Unlike other sex and race-specific lung cancer, the correlograms of nonwhite female case between the two different time periods were very dissimilar. The latter period had higher spatial autocorrelation indices than the earlier period. Significantly positive autocorrelation in the latter one was generated at the lags 1, 2, and 3, whereas negative autocorrelation was shown at the lags 6, 7, and 8.

5.3.2 Stomach Cancer

Like lung cancer case, stomach cancer had positive autocorrelation, which existed among physically contiguous parishes (Figure 5.11). For 1953-1987, high positive autocorrelation occurred at lags 1 and 2 and negative autocorrelation occurred at lags 5 and 6. Generally, stomach's correlograms showed a rapidly declining irregular V-shape through distance.

The curve of correlograms in 1953-1977 was nearly the same as that in 1953-1987 (Figures 5.12 and 5.13). However, the curve of correlograms increased from

lag 1 to lag 2 and then dropped considerably through lag 6 in 1978-1987. This could be interpreted that more prominent autocorrelation existed in the later period, suggesting large amounts of second order neighbors.

For 1953-1987, white male stomach cancer had significant positive autocorrelation at lags 1 and 2 and negative autocorrelation at lag 5, whereas white female stomach cancer showed no autocorrelation (Figures 5.18 and 5.19). Generally, the correlograms of cancer mortality rates for white males showed irregular mountain-shape and those of white females generated several small V-shapes.

Comparing the two time periods for white male stomach cancer, the positive autocorrelation indices of cancer mortality rates for white males in 1953-1977 were high at lag 2, whereas those in 1978-1987 were at lag 4. In case of negative autocorrelation indices for white male stomach cancer, the earlier period was high at lags 4 and 5, whereas the latter period was high at lags 6 and 7. In other words, at lag 2, the high value of correlogram in 1953-1977 dropped below significant level in 1978-1987. Furthermore, at lag 4, the earlier period showed significant negative autocorrelation, whereas the latter period had a high positive autocorrelation value. At this distance from a given parish, it is likely that parishes in clusters of dissimilar values were present in 1953-1977, and that parishes in clusters of similar values were present in 1978-1987. This indicated the presence of rather extended small clusters in the southern portion of the state over time. Stomach cancer among white females, shown in Figure 5.19, provided insignificant autocorrelation value, showing random distribution of cancer mortality.

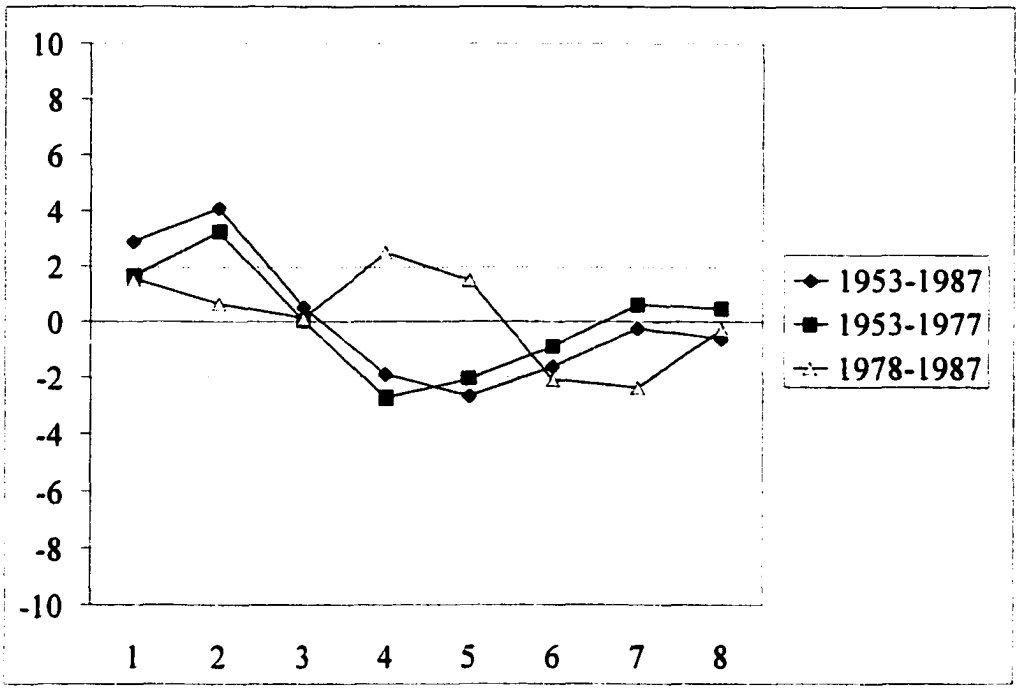

Figure 5.18 Correlograms for Stomach Cancer (White Males)

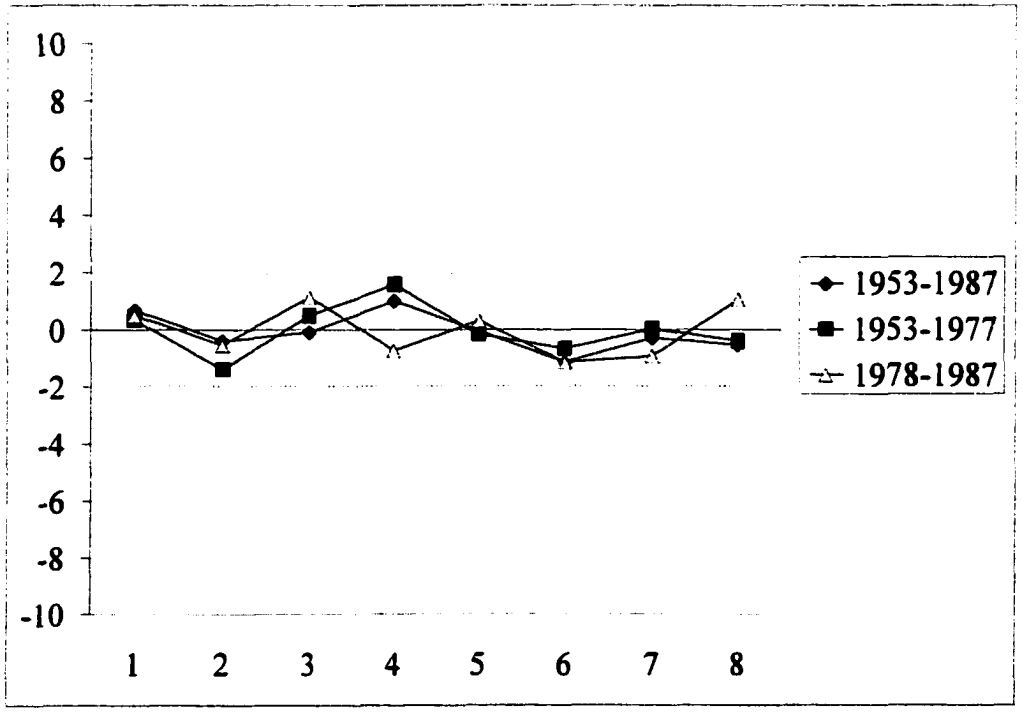

Figure 5.19 Correlograms for Stomach Cancer (White Females)

Figures 5.20 and 5.21 show the correlograms of stomach cancer among nonwhites from 1953 to1987. Standard normal deviate value in male stomach cancer ranged from -4.09 to 7.09, whereas that in females ranged from -2.49 to 4.13. Stomach cancer for nonwhite males had the second highest positive among the 16 cancer sites and the autocorrelation indices had negatively or positively significant level, except for lag 5. The forms of correlograms between both sexes for nonwhite stomach cancers (from 1953 to 1987) showed relatively similar form (except for lag 5). Their curves maintained the similar form of decline with distance, and then increased. But males were more consistent through time than females.

During the two time periods, the earlier period's correlogram for nonwhite males showed declining V shape and the latter period's depicted a roughly declining curve (except for lag 1). This indicated a small difference in geographic distributions of high cancer rates through time. Centers of the neighboring parishes (with high cancer mortality rates) were west of lower Mississippi River in the earlier period whereas in the latter period they were in lower southern Louisiana. Nonwhite female stomach cancer for 1953-1987 showed significantly positive autocorrelation at lag 1 and negative autocorrelation at lag 6. Specifically the two periods had no autocorrelation at each lag and showed generally stable forms of correlograms.

5.3.3 Breast Cancer

Among the five cancer sites, the value of positive spatial autocorrelation for breast cancer remained significantly high at the first or second spatial lag (Figure 5.11). In other words, there was significant autocorrelation among the second order neighbors as well as among the first order neighbors. It might only explain the presence of rather

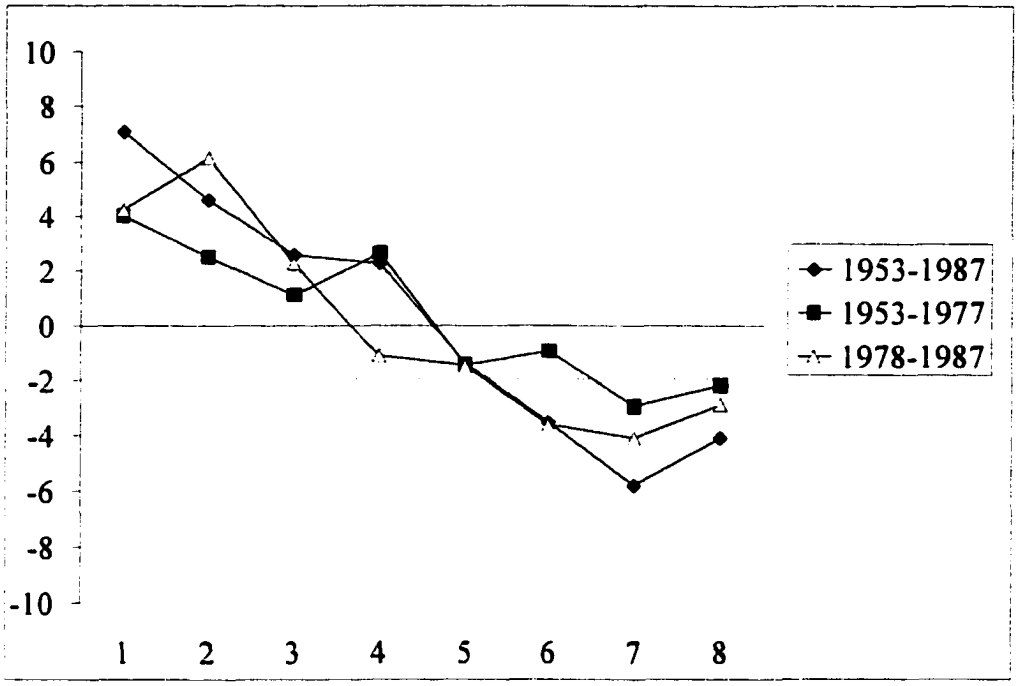

Figure 5.20 Correlograms for Stomach Cancer (Nonwhite Males)

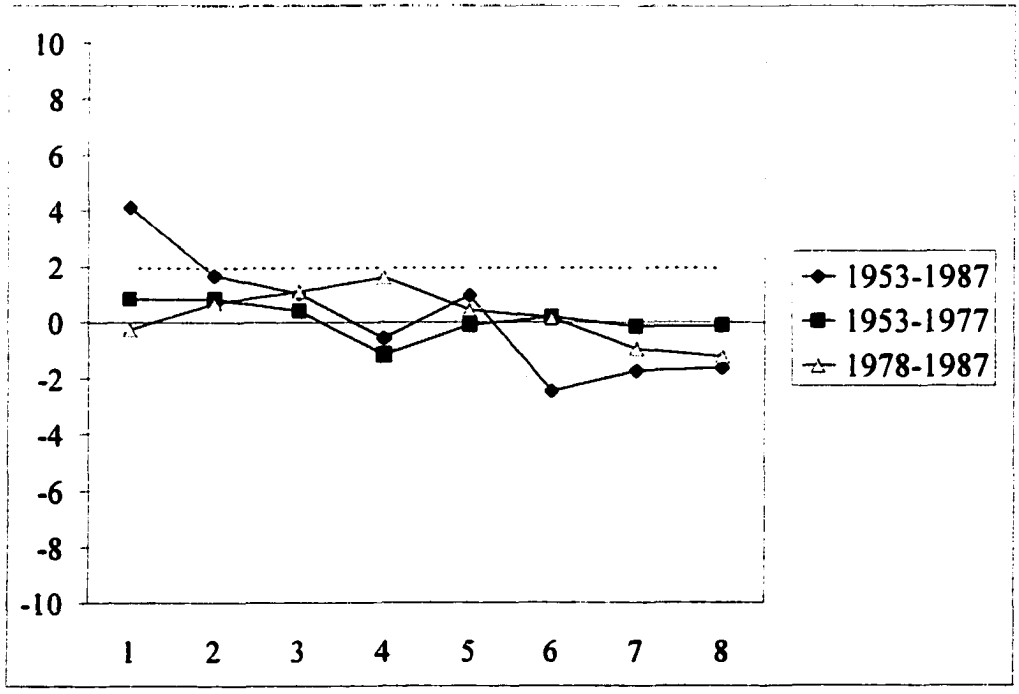

Figure 5.21 Correlograms for Stomach Cancer (Nonwhite Females)

large clusters of parishes with similar breast cancer rates. Like previously mentioned cancers, the correlograms had very similar declining and increasing curves in the two different time periods (Figures 5.12 and 5.13), with a uniform spatial distribution or clustering of breast cancer rates.

In case of race-specific breast cancer, autocorrelation values were insignificant, indicating a random relationship between parishes. However, only nonwhite breast cancer for 1978-1987 had significantly positive autocorrelation at lags 1 and 2 and negative autocorrelation at lags 6 and 8 (Figures 5.22 and 5.23). This indicated the presence of rather large clusters of parishes with similar (high) breast cancer rates in southern Louisiana and with dissimilar (low) rates in northern Louisiana.

5.3.4 Prostate Cancer

As shown in Figure 5.11, the spatial autocorrelation computed for prostate cancer mortality rates was below the statistically significant level, except for the unexpectedly highly negative vale at lag 8 for 1953-1987. No definite pattern of correlograms could be found, but the shape irregularly fluctuated, rapidly dropping at lag 8. In the two time periods (Figures 5.12 and 5.13), all autocorrelation values for the earlier period were insignificant, indicating a random relationship between parishes. Similar to that of 1953-1987, the general form of the correlograms throughout the latter period was quite stable, with a high negative autocorrelation point at lag 8.

Prostate cancer among white males had only a negative autocorrelation point at lag 8 in 1978-1987 (Figure 5.24), whereas nonwhite males had a negative autocorrelation point at lag 1 in 1953-1977 or at lag 2 in 1978-1987 (Figure 5.25). Through time, the correlogram's form among white males was changed from stable or a

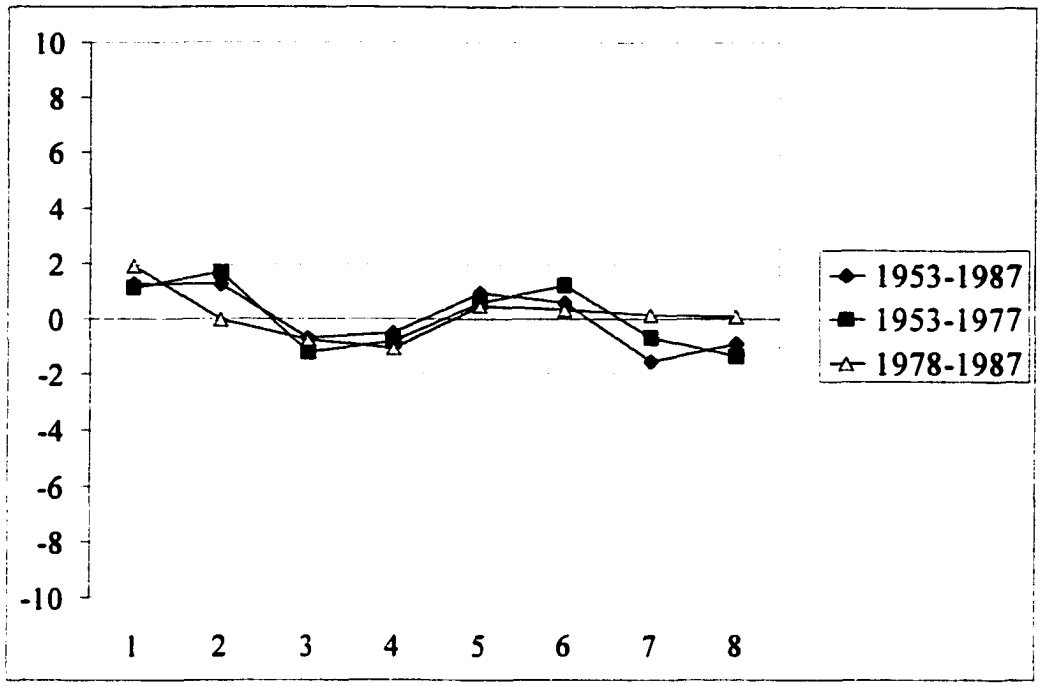

Figure 5.22 Correlograms for Breast Cancer (White Females)

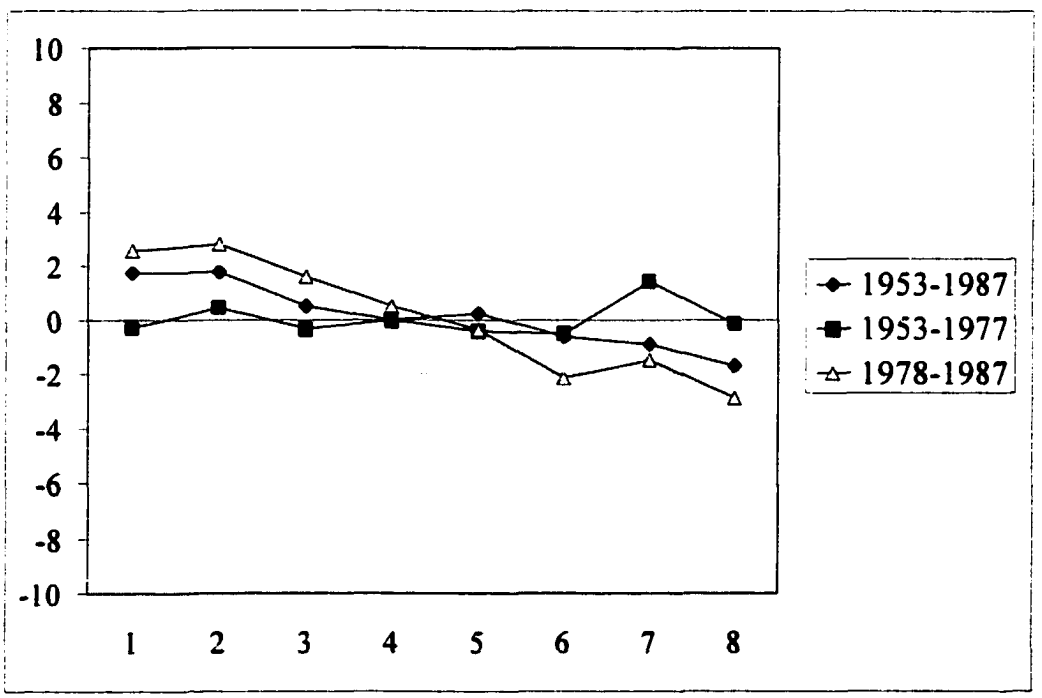

Figure 5.23 Correlograms for Breast Cancer (Nonwhite Females)

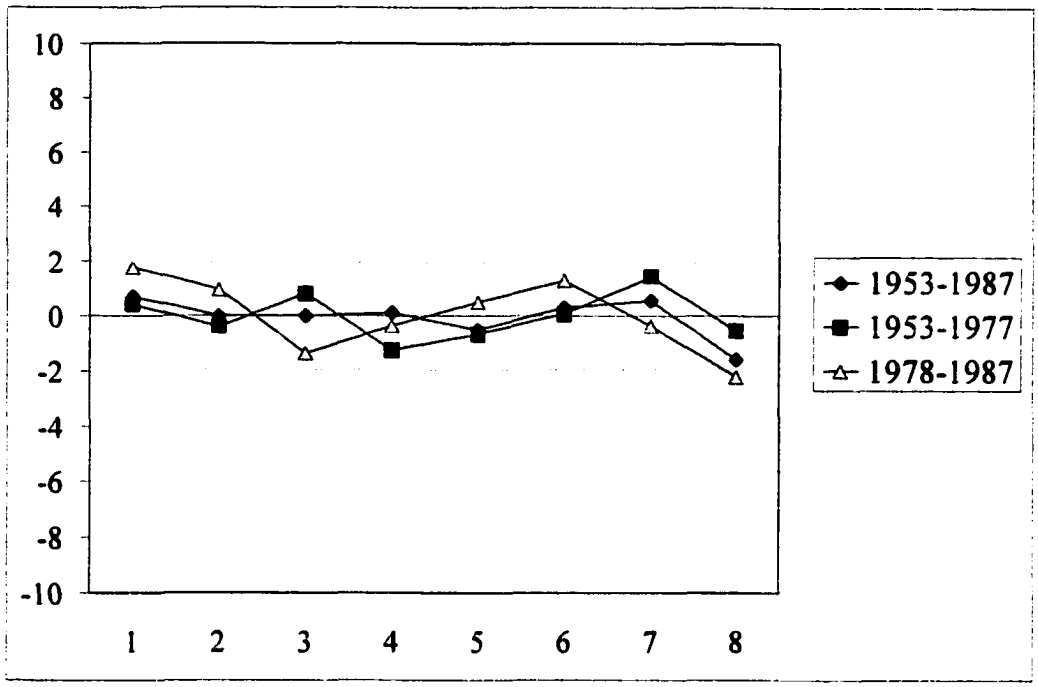

Figure 5.24 Correlograms for Prostate Cancer (White Males)

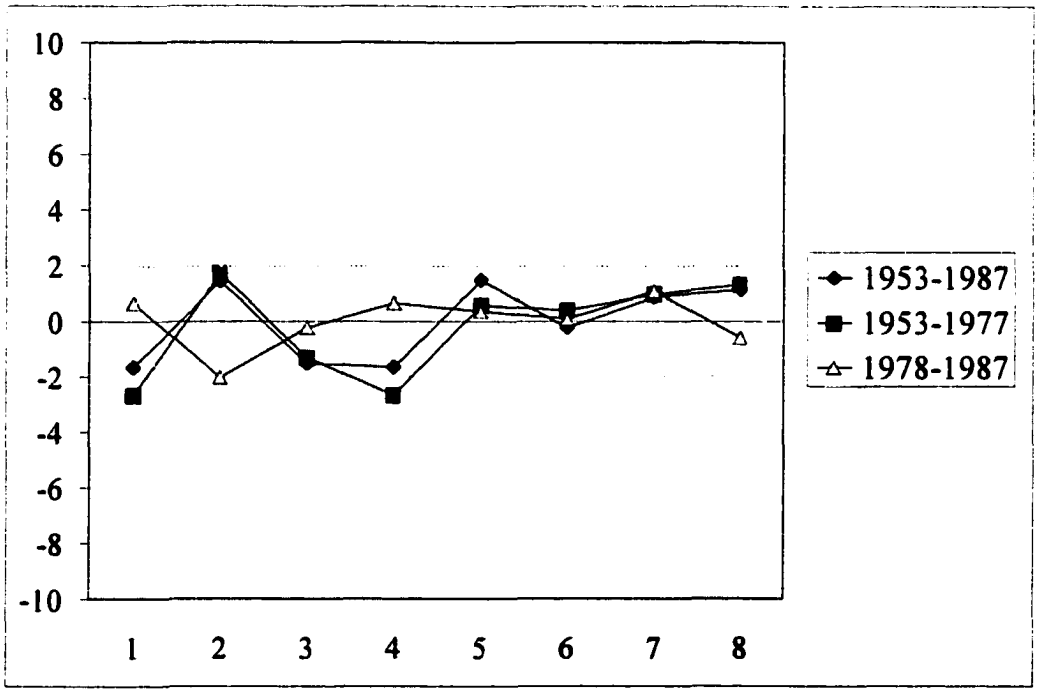

Figure 5.25 Correlograms for Prostate Cancer (Nonwhite Males)

little fluctuated V shape to wide V shape. In case of nonwhite males, the correlogram shape at the first four lags during two different periods looked totally opposite. This suggested clearly that the distribution of prostate cancer mortality rates among white males had showed small change and that the distribution among nonwhite males had generated big change. Unlike other cancer sites, prostate cancer had only high negative autocorrelation. White males had negative autocorrelation at lag 8, whereas nonwhite males had that at lag 1 or 2.

5.3.5 Colorectal Cancer

For 1953-1987, the spatial autocorrelation indices for colorectal cancer were generally very low and clearly exhibited no neighborhood effects, except for significant autocorrelation at spatial lag 1 (Figure 5.11). The pattern of correlograms showed a rapidly dropping curve at lags 1 and 2 and then smoothly declining one, indicating a random relationship between parishes. In case of the comparison of the two time periods (Figures 5.12 and 5.13), the general form of the correlograms was a V shape in 1953-1977 and mountainous shape with totally opposite peak (with only high positive autocorrelation point at lag 2) in1978-1987. This explained that colorectal cancer mortality rates fluctuated from one parish to another and their distribution was quite dispersed, with clustering in parishes neighboring New Orleans.

Comparing colorectal cancer rates by sex and race, shown in Figures 5.26 and 5.27, the spatial correlograms for white females and nonwhite males provided evidence of no autocorrelation. Among white males (Figure 5.26) and nonwhite females (Figure 5.29), the positive spatial autocorrelation usually existed at lag 1 or 2. For 1953-1987, colorectal cancer of white males and nonwhite females had similar positive

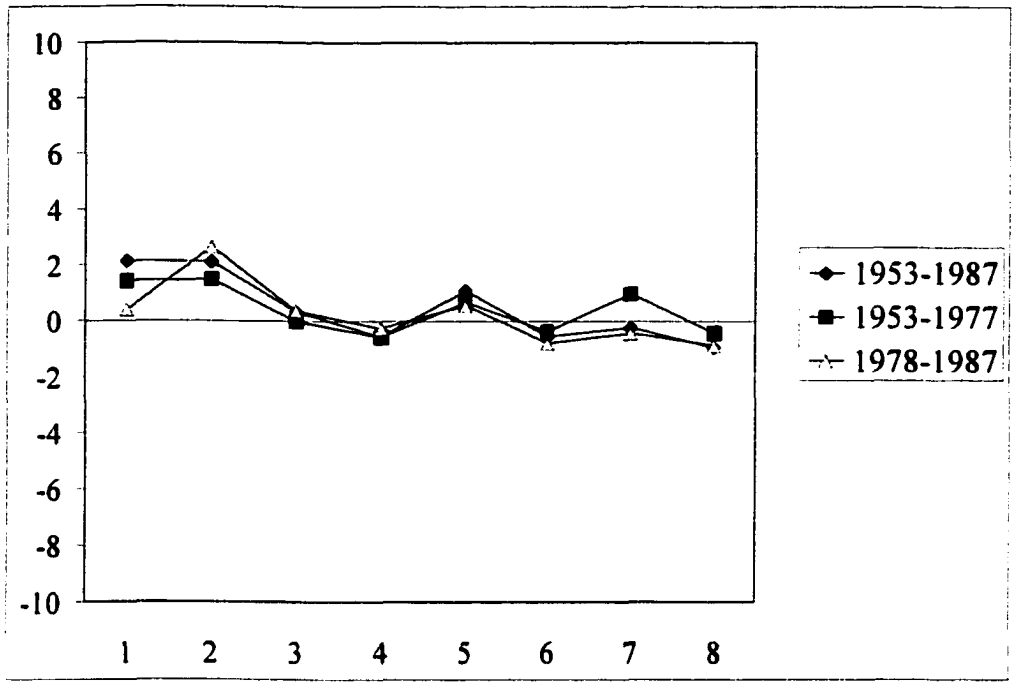

Figure 5.26 Correlograms for Colorectal Cancer (White Males)

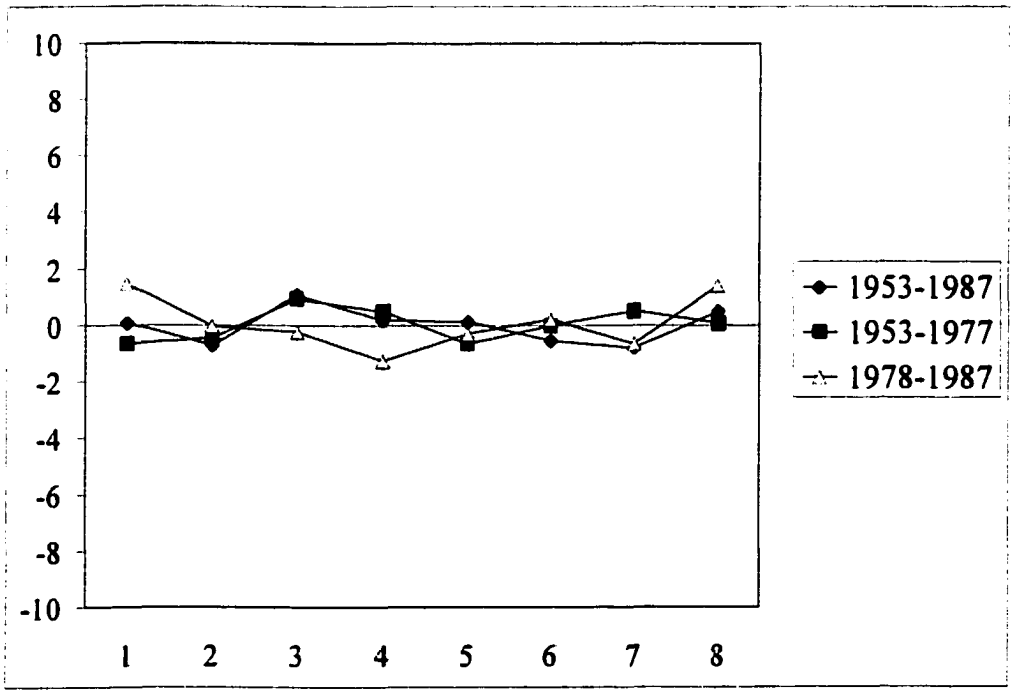

Figure 5.27 Correlograms for Colorectal Cancer (White Females)

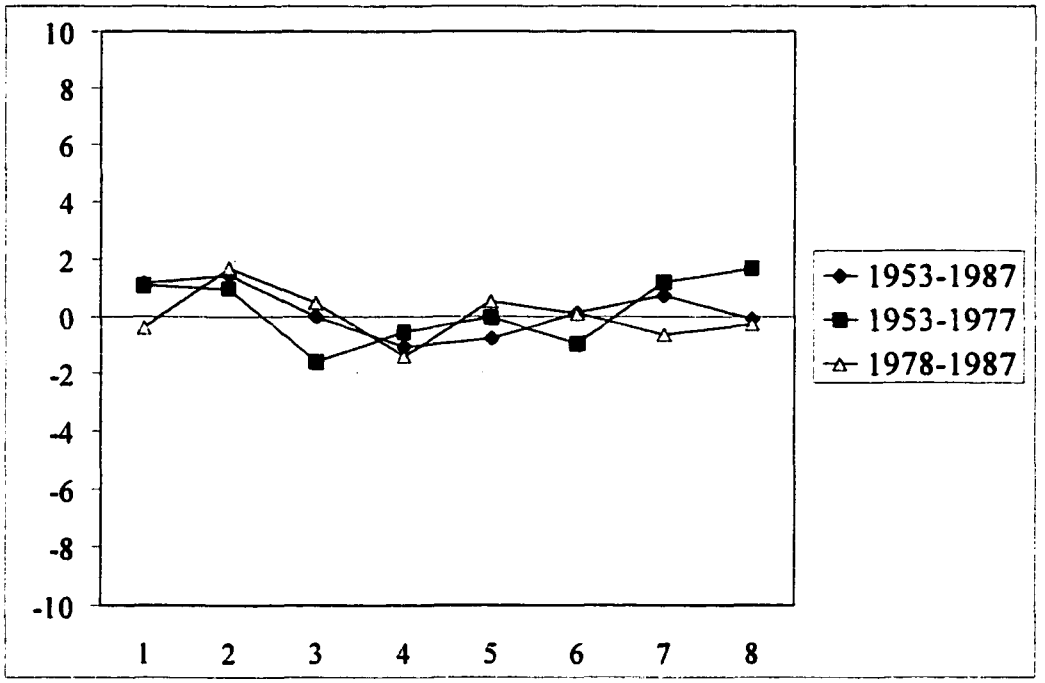

Figure 5.28 Correlograms for Colorectal Cancer (Nonwhite Males)

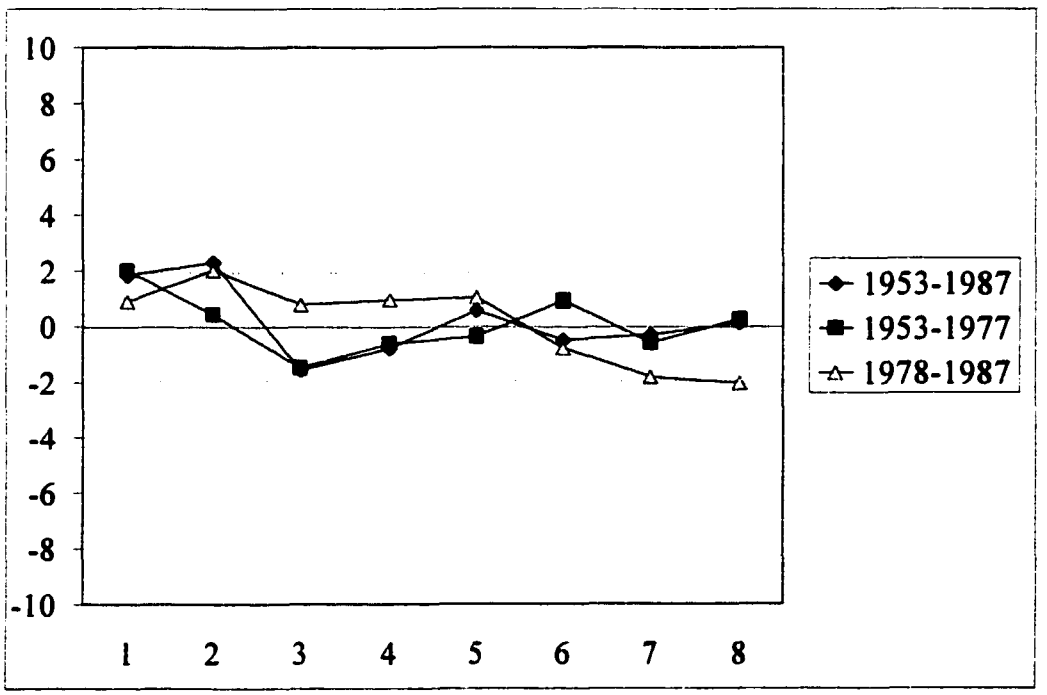

Figure 5.29 Correlograms for Colorectal Cancer (Nonwhite Females)

autocorrelation at lags 1 and 2 and parts of the correlograms showed a horizontal line of curve. It corresponded to parishes in the end of southeastern Louisiana with similar colorectal cancer mortality rates.

In particular, the latter period (1978-1987) generated similar form of correlograms, which dropped below significance at lag 1 and increased above significance at lag 2, except for white males. It indicated generally that a large amount of second order neighbors had similar colorectal cancer rates.

The findings of this study, as summarized above, showed quite a wide variation of spatial autocorrelation among cancer sites through time. Spatial correlograms described spatial autocorrelation or spatial clustering at different spatial lags. They provided an effective means of comparing the spatial-temporal patterns of cancer mortality rates among parishes. Generally, cancers of lung and stomach exhibited the strongest degree of positive or negative spatial autocorrelation, indicating that cancer mortality rates among parishes showed a significant amount of spatial differentiation. The spatial correlograms for cancers of breast, colorectum, and prostate have low spatial autocorrelations, indicating weak spatial differentiation exhibited by these cancer sites.

As in previous studies (Glick 1977; Kennedy 1988), male lung cancer as well as stomach cancer among both sexes exhibited high spatial autocorrelation. However, female lung cancer did not correspond to previous studies, which had low spatial autocorrelation. In particular, of all the correlograms, lung cancer's correlograms were the most drastic and distinctive. Not only were the spatial autocorrelations high, but also the correlograms changed considerably through time. As time passed, the

distribution of high mortality rates for males tended to be distributed widely and randomly whereas that for females tended to be clustered in a small region.

As earlier studies suggested (Glick 1977, 1979; Kennedy 1988; Lam et al. 1996), the results from this correlogram analysis provided useful insights into the specification of disease models for possible causal factors or more accurate forecasting. If there is little variation in the form of correlograms among both sexes of whites and nonwhites, then a uniform spatial distribution model may be applied subsequently to explain the occurrence of cancer. If the correlograms differ substantially among them, spatial models with different parameters and factors must be specified to produce more accurate explanation.

The spatial and temporal resolution of cancer data as well as site, sex, and race-specific cancer rates affects the interpretation of data, and therefore, the correlograms must be interpreted with caution. Because of the problem of scale dependency, the choice of a hypothesis-testing assumption (normal versus randomization) and the definition of spatial adjacency matrix (binary or continuous form) might affect the results. Also, the forms of the correlograms may be influenced by the settlement pattern or urban structure of the region.

In summary, as time progresses, the distribution of high cancer mortality rates has become more widely and randomly distributed. It is safe to say that the autocorrelation and correlogram analyses indeed suggested hypotheses (such as high degree of spatial clusterings for lung and stomach cancer) that need to be explored further. Further progress in understanding the geographic distribution of mortality rates will come from additional testing of existing and newly developed hypotheses. One

approach is to examine cancer mortality statistics and environmental factors in areas having high mortality and in those having low mortality. Research on the effect of environmental variables on cancer incidence, mortality, and survival in Louisiana is much needed. Therefore, the relationship between cancer mortality patterns and environmental factors in Louisiana is investigated in the following sections.

5.4 Factor Analysis of Enviromental Variables

Since many environment variables in this analysis are a highly correlated (Appendix S), a factor analysis was performed on the initial 24 variables to find a more parsimonious set and to estimate the characteristics of the selected environmental variables. Their results are briefly summarized below.

Factor analysis of the 24 variables using the principal axis components option resulted in six factors. Table 5.10 presents the percentage total variance explained by each factor. The six factors retained and rotated explain 64.7% of total variance among the variables. Factor 1 explains 16.9% of the total variation of the data set. Factors 2, 3, 4, 5, and 6 explain 15.1%, 13.6%, 6.8%, 6.5%, and 6.0%, respectively. As can be seen in Table 5.10, persons below poverty level (0.97), persons who employed in chemical industry (0.97), and wetlands (0.95) had the highest communality among 24 variables. The high communality values of these variables could be easily expected because of the dominance of below the poverty level, the largest single employer in the Louisiana as well as the U.S manufacturing sector, and huge wetlands state (40 % of the U.S.).

Table 5.11 shows the results of the varimax rotation of the six factors to make the factors more interpretable. Factor 1 represents a general factor with four of the

Table 5.10 Total Variance Explained for Factor Analysis of Environmental Variables in Louisiana for 1980-1989 (by six factors)

		Initial Eigenvalues				Extraction Sums Sq.Loadings			Rotation Sums Sq. Loadings		
Variables	Communalitites	Factor	Total(%)	Variance	Cumulative	Total(%)	Variance	Cumulative	Total(%)	Variance	Cumulative
NWP	.622	1	7.351	30.627	30.627	7.126	29.691	29.691	4.051	16.881	16.881
PD	.606	2	2.954	12.307	42.934	2.638	10.994	40.685	3.613	15.054	31.935
PCPI	.831	3	2.410	10.042	52.977	2.054	8.557	49.241	3.257	13.571	45.505
PBPL	.965	4	2.303	9.594	62.571	1.991	8.295	57.536	1.623	6.765	52.270
EDS	.793	5	1.480	6.165	68.736	1.065	4.439	61.976	1.554	6.476	58.746
AGR_EM	.765	6	1.219	5.078	73.814	.663	2.762	64.737	1.438	5.991	64.737
MIN_EM	.519										
CON_EM	.512										
MAN_EM	.506										
CHM_EM	.964										
TRA_EM	.842										
HS_EM	.316										
EDS_EM	.259										
TTRI	.783										
TCARC	.621										
H_WS	.827										
S_WS	.141										
NPL	.613										
AGR_CH	.697										
WET	.947										
UR_POP	.654										
D_WAT_M	.813										
D_WAT_S	.605										
OZONE	.348										

Extraction Method: Principal Axis Factoring
Rotation Method: Varimax

(Table 5.10 Continued)

Variables: Normalized Variables – See Environmental Data in section 3.2.2

AGR_CH	:Pesticide (Agricultural chemicals)
AGR_EM	:Agriculture (Persons employed)
CHM_EM	:Chemical manufacturing (Persons employed)
CON_EM	:Construction (Persons employed)
D_W_M	:Mississippi River (Drinking)
D_W_S	:Surface water (Drinking)
EDS	:Education status (High school graduates)
EDS_ EM	:Educational services (Persons employed)
H_S_EM	:Health services (Persons employed)
H_WS	:Hazardous waste sites
MAN_EM	:Manufacturing (Persons employed)
MIN_EM	:Mining (Persons employed)
NPL	:National Priorities List (Waste sites)
NWP	:Nonwhite population
OZONE	:Ozone exceedances (Air quality)
PBPL	:Persons below poverty
PCPI	:Per capita income
PD	:Population density
S_WS	:Solid waste sites
TCARC	:Total Carcinogenic TRI
TRA_EM	:Transportation (Persons employed)
TTRI	:Total Toxics Release Inventory (TRI)
UR_POP	:Urban population
WET	:Wetlands

Table 5.11 Rotated Factor Matrix for Environmental Variables (Louisiana, 1980s)

	Factor 1	Factor 2	Factor 3	Factor 4	Factor 5	Factor 6
Wetlands	0.863	0.166	0.229	-0.218	-0.189	-0.199
Mississippi River (Drinking)	0.818	0.130	0.117	0.253	0.181	-0.128
Surface water (Drinking)	0.766	-0.030	0.090	0.020	-0.017	0.089
Transportation (Persons employed)	0.718	0.159	0.349	0.027	-0.420	-0.049
Urban population	0.533	0.054	0.152	0.323	-0.075	0.483
Total Toxics Release Inventory (TRI)	0.529	0.500	0.108	0.321	0.242	-0.282
Per capita income	0.526	0.289	0.360	0.307	-0.234	0.438
Chemical manufacturing (Persons employed)	0.210	0.909	0.112	-0.162	0.236	0.012
National Priorities List (Waste sites)	-0.130	0.762	0.068	0.021	-0.028	0.092
Total Carcinogenic TRI	0.158	0.745	0.089	0.124	0.136	0.084
Hazardous waste sites	0.445	0.612	0.100	0.454	-0.042	0.214
Construction (Persons employed)	0.194	0.547	0.271	-0.198	-0.162	-0.191
Ozone exceedances (Air quality)	-0.028	0.447	0.050	0.064	0.052	0.371
Persons below poverty	-0.260	-0.318	-0.824	-0.104	0.320	-0.064
Agriculture (Persons employed)	-0.199	-0.147	-0.813	-0.095	0.080	-0.159
Pesticide (Agricultural chemicals)	-0.152	0.019	-0.806	-0.056	-0.135	-0.052
Nonwhite population	0.022	-0.051	-0.555	0.042	0.521	0.192
Education status (High school graduates)	0.180	0.237	0.548	0.406	0.051	0.486
Manufacturing (Persons employed)	0.127	0.378	0.405	-0.155	0.371	-0.145
Population density	0.402	0.056	0.030	0.660	-0.070	0.013
Health services (Persons employed)	-0.124	-0.078	0.140	0.463	0.073	0.234
Mining (Persons employed)	0.107	-0.190	0.087	-0.171	-0.658	-0.045
Solid waste sites	-0.033	0.034	0.074	-0.144	0.293	0.161
Educational services (Persons employed)	-0.055	-0.013	0.010	0.062	0.152	0.478

Variables: Normalized Variables – See Environmental Data in section 3.2.2
Extraction Method: Principal Axis Factoring
Rotation Method: Varimax

twenty-four variables loading at 0.65 or higher. Physical environmental variables related to water (wetlands, Mississippi River or surface water for drinking water) loaded high on factor 1. Since wetlands had the highest variable loading on factor 1 it was chosen to represent this factor. High scores for this factor were dominantly in southern coastal Louisiana.

Four variables loaded very highly onto factor 2: chemical manufacturing (0.91), NPL waste sites (0.76), total carcinogenic TRI (0.75), and hazardous waste sites (0.61). This factor seems to be best represented by the chemical pollution variables. High scores for this factor were in the industrial corridor along the Mississippi River and some parishes of southwestern Louisiana.

Factor 3 showed high negative factor loadings of persons below poverty (-0.82), pesticide use in agriculture (-0.81), persons employed in agriculture (-0.81), and nonwhite population (-0.56). Negatively high scores for this factor were dominantly in northeastern Louisiana. These parishes had relatively low-educated and low income inhabitants who are engaged in agriculture. One of the major agricultural policy in the U.S. was to support low farm income via government subsidies during the study period. For example, Louisiana rice producers were supported through the government deficiency payment (Knutson et al. 1998). Parishes which had (negatively) high scores had lower per capita income, higher nonwhite population, lower rates of high school graduation, and more rural areas than other parishes.

Population density (0.67) loaded highest on Factor 4. Health services did not have high loading value (0.46) but it loaded second highest on this factor. Wetlands that are not highly populated loaded the first negatively high loading on this factor

(-0.22). This factor could represent a measure of resources required for high population. Factors 5 and 6 were not considered a pervasive influence in this analysis because of the low degree of variance they explain.

According to the results of factor analysis, the environmental variables in this research could be grouped into four categories: physical environment, chemical pollution, socio economic status, and demographic resource.

5.5 Multiple Regression Analysis of Cancer Mortality Rates and Environmental Factors

This subchapter focuses on mortality data within Louisiana to generate and test hypotheses for cancer etiology, stressing both the limitations and advantages of this approach. A series of correlation and regression studies linking cancer mortality rates with demographic, socioeconomic, industrial, and environmental data at the parish level from 1980 to1989 were conducted. The results of multiple correlation and regression analyses are summarized below.

Appendix S shows the Pearson's correlation coefficients among all dependent and independent variables. Many of the independent variables were found to be rather highly correlated, such as persons employed in agriculture and pesticide (agricultural chemicals) (0.78), wetlands and persons employed in transportation (0.77), total carcinogenic TRI and persons employed in chemical manufacturing (0.71), and per capita income and education status (0.71). These high correlations may not be critically problematic. However, a correlation coefficient between two explanatory variables greater than 0.8 or 0.9 indicates a strong linear association and a potentially harmful collinear relationship (Griffiths et al. 1993).

Multiple stepwise regression analysis was performed, and the results of the stepwise regression analysis for combined cancer mortality rates are presented in Table 5.12 and Appendix T. It indicated that there were three predictor variables that could significantly contribute one percent or more to the explained variance in support for high cancer mortality rates. Urban population was the first variable to enter the model. It explained 25.6% of the variation in support for cancer mortality rates. Persons employed in health service were the second variable to enter the model, adding 10.7% to the explanatory power. Persons employed in education service were the third variable and added another 4.2%. The combined model of three variables explained 40.5% of the total variation of the cancer mortality rates. The multiple regression equation estimated above suggested several findings. Urban population appeared to be the best predicator of cancer mortality rates, at least among the variables included in this study. Many studies indicated that urban population was highly related to cancer mortality rates. A possible explanation of the observed results was that urban residents have some environmental characteristics to increase cancer death rates. Persons employed in health service and education service also appeared to be important variables and showed a negative regression coefficient. People who worked for health and education services had lower cancer mortality rates than persons employed in any other occupation because they had more chance to decrease or prevent the death rates by the awareness of disease and economic status,

Site, sex, race-specific cancer mortality rates as the dependent variables were studied using multiple regression analysis (Table 5.13), to gain more insights into the association between cancer and the environment.

Table 5.12 Stepwise Regression Analysis of Cancer Mortality Rates for the Cancer Sites

Dep. / Ind.	All Sites	Breast	Colorectum	Lung	Prostate	Stomach
AGR_EM						
CHM_EM						
CON_EM						
EDS						
EDS_EM	(-)					
H_W_S						
H_S_EM	(-)					
MAN_EM						
MIN_EM						
D_W_M			(+)	(+)		
NPL						
NWP					(+)	(+)
OZONE						
PCPI						
PBPL				(-)		
AGR_CH		(-)	(+)	(+)		
PD						
S_W_S						
D_W_S						
TCARC						
TTRI						
TRA_EM				(+)		
UR_POP	(+)	(+)	(+)			
WET					(-)	(+)
R^2	.405	.261	.316	.527	.375	.273

See Table 5.10 for the independent variables defined. See Appendix 5.12 for the regression output in detail.

Table 5.13 Stepwise Regression Analysis of Cancer Mortality Rates for the Cancer Sites (by Sexes and Races)

Ind.	**Breast**		**Colorectum**				**Lung**				**Prostate**		**Stomach**			
Dep.	wf	nwf	wm	wf	nwm	nwf	wm	wf	nwm	nwf	wm	nwm	wm	wf	nwm	nwf
AGR_EM							(+)									
CHM_EM															(+)	
CON_EM				(-)				(+)		(+)						
EDS	(+)				(-)				(-)		(+)					
EDS_EM																
H_W_S																
H_S_EM														(-)		
MAN_EM	(+)						(+)									
MIN_EM		(+)						(+)							(+)	
D_W_M			(+)													
NPL																
NWP																
OZONE																
PCPI				(+)												
PBPL									(-)		(+)					
AGR_CH		(-)		(+)												
PD																
S_W_S																
D_W_S								(-)								
TCARC																
TTRI						(-)								(+)		
TRA_EM							(+)					(-)	(+)			
UR_POP					(+)			(+)	(+)	(+)		(+)				
WET				(+)		(+)									(+)	(+)
R^2	.26	.23	.10	.33	.36	.23	.32	.42	.31	.33	.29	.14	.10	.17	.52	.12

See Table 5.10 for the independent variables defined. See Appendices T for regression outputs in detail.
wm: white males, wf: white females, nwm: nonwhite males, nwf: nonwhite females

5.5.1 Lung Cancer

In Table 5.12, a stepwise multiple regression analysis showed that transportation, agricultural chemicals, and Mississippi Rivers as drinking waters had a positive relationship with cancer mortality rates for lung cancer whereas persons below poverty had a negative relationship with them. Persons employed in transportation were the first variable to enter the model. This value was very high, compared with other variables (Appendix T). It indicated that the distributions of the employees in transportation were very similar to those of lung cancer mortality rates. *The 1987 Census of Transportation* recorded establishments within three categories: (1) motor freight transportation and warehousing - includes usually general freight carriers, moving companies, garbage freight carriers, moving companies, garbage and trash collection, dump trucking, and courier services; (2) water transportation - includes freight transportation, passenger transportation, and water transportation service; (3) transportation service - includes travel agencies, freight shipping service, and car rentals (Goins and Caldwell 1995). During 1985-1989, twenty parishes (along the Gulf Coast and the Mississippi River, where maritime occupations are numerous) were above the state's labor average (Goins and Caldwell 1995). Unlike a previous multiple regression study in which the largest industry with the highest lung cancer rates was the chemical industry (Blot and Fraumeni 1976), transportation had extremely closer relationship with lung cancer mortality rates than those in other industries, possibly because the previous study considered only 27 Louisiana parishes.

Agricultural chemicals were the second variable to the model. Louisiana is the nation's second largest producer of agricultural chemicals (after Florida). Agricultural chemicals used primarily as pesticides and herbicides and production in this area includes nitrogen gertilizers and phoshate fertilizers (Scott 1993).

Mississippi River is the only source of drinking water for many industrial corridor residents. The Mississippi River industrial corridor has more than 128 industrial facilities reporting TRI emissions to the EPA (LDEQ 1991). Many of the facilities release chemical substances which have known or suspected effects on the health of human or aquatic life or wildlife. Many studies have suggested a relationship between chemical exposure and health effects, but they are not able to make a direct correlation between health effects and the amount of chemicals released into the environment

With regard to white males (Table 5.13), the results of a stepwise regression analysis indicated a positive relationship between the cancer and persons employed in agriculture, in addition to persons employed in transportation and Mississippi River. Parishes which have many employees in agriculture release many agricultural chemicals. The geographic distributions of two variables were very similar to those of lung cancer mortality rates among white males, which were dominantly distributed in southeastern and northeastern Louisiana. Lung cancer mortality rates among white females indicated a positive relationship for urban population and persons who employed in mining and construction and a negative relationship for surface water.

The analysis of stepwise regression identified three variables [persons below poverty (-), urban population (+), and education status (-)] for nonwhite males and two

variables [urban population (+) and persons employed in construction (+)] for nonwhite females in the model. Persons below poverty negatively affected lung cancer mortality rates for total and nonwhite males. Parishes which had high income (by industrialization and urbanization in Louisiana) showed high cancer mortality for lung cancer. However, some studies (Chen et al. 1944 and Groves et al. 1996) suggested that the high death rates in South Louisiana are not due to high incidence rates but poor cancer prognosis and high poverty.

In particular, occupational variables in the regression specification of lung cancer mortality rates were more important than those variables in the regression equations of any other cancer mortality rates. It explained that some risk factors in lung cancer mortality rates include exposure to certain industrial substances, such as arsenic, some organic chemicals, and radon and asbestos.

5.5.2 Breast Cancer

A stepwise multiple regression analysis indicated that urban population had a positive effect on breast cancer mortality rates; on the other hand, agricultural chemicals negatively affected breast cancer mortality rates. Do these results indicate that the more agricultural chemicals exist, the less breast cancer mortality rates occur? It did not imply that. A possible explanation of the observed results was that persons living in the parishes which had the high emission of agricultural chemicals tended to work for agriculture or fertilizer industry, and had low income.

The multiple regression equation among white females suggested that education status (+) appeared to the best predictor and persons employed in manufacturing (+) was the second variable to enter the model. Persons employed in education service were

expected to have a positive effect on high breast cancer rates since women who work for education service have high education and socioeconomic status, never having children or having the first live birth at a late age. Furthermore, women prefer to have a job in education service and have more jobs in that occupation than other occupations. Therefore, persons employed in these jobs would be able to have more chance to get high breast cancer death rates. In case of nonwhite females, persons employed in mining (+) and agricultural chemicals (-) were the first and second variable to enter the model, respectively. Therefore, the cancer mortality rates among white females were more closely related to education status than nonwhite females.

5.5.3 Colorectal Cancer

In colorectal cancer, the variables in the order in which they entered the stepwise regression model included Mississippi River as drinking water, agricultural chemicals, and urban population. Mississippi River as drinking water was identified as a fitted variable for the model of colorectal cancer mortality rates among white males whereas four variables [agricultural chemicals (+), per capita income (+), persons employed in construction (-), and wetlands (+)] were identified as selected variables for that among white females (Table5.13). Previous studies (Gottlieb et al. 1981, 1982a, and 1982b) showed an increased risk for cancer of the rectum in those who consume Mississippi River water compared with those who drink groundwater. As mentioned before, Mississippi River is the only source of drinking water for industrial corridor residents, and conventional water treatment methods do not remove herbicides from drinking water (Institute for Environmental Issues and Policy Assessment 1994).

In case of the nonwhite males, there were two predictor variables in this model: urban population (+) and education status (-). Unlike breast cancer case, education status negatively affected colorectal cancer mortality rates among nonwhite males. For nonwhite females, there were two predictor variables in this model, wetlands and total toxics releases. Wetlands had a positive impact on the cancer death rates. Especially, the regression equation with the total toxics releases indicated a negative association between them. These unexpected signs were partly due to the lack of data availability of large scale. In addition, the regression equation with stepwise selection reduced the value of R-square (Appendix T). Therefore, it is important to find out how this variable is associated with the cancer death rates, by adding more variables or accomplishing different analysis methods.

5.5.4 Prostate Cancer

A model for the prostate cancer mortality rates showed that nonwhite population (+) and wetlands (-) explained the total variation of the death rates by 37.5 %. Nonwhite population variable appeared to be important in this analysis. American Cancer Society (1998) reported that prostate cancer mortality rates are more than two times higher for African-American men than white men. Also, the geographic distributions of nonwhite population and prostate cancer mortality rates showed a very similar pattern. They were distributed in northeastern Louisiana and upper Mississippi River, not southern Louisiana. The wetland (which was dominantly distributed in coastal Louisiana) variable negatively affected prostate cancer mortality rates. The geographically opposite distributions between wetlands and prostate cancer rates might be attributed to this negative relationship between them.

Prostate cancer mortality rates among white males showed two predictor variables to explain the models: person below poverty (+) and education status (+). It was difficult to conceptualize the model because these two variables have socio-economically opposite characteristics. For nonwhite males, employees in transportation had a negative effect to the death rates of prostate cancer and urban population had a positive impact on them. Like wetlands, the distributions of employees in transportation showed opposite distributions of high cancer mortality rates.

5.5.5 Stomach Cancer

Wetlands appeared to be the best predictor of stomach cancer mortality rates for total and nonwhites. Wetlands are important to wildlife and for maintaining water quality. Southern coastal Louisiana which contains about 40% of the U.S. wetlands showed significantly high stomach cancer rates for total and nonwhites. A better understanding of the characteristics of not only wetlands but also areas that include many wetlands is needed. Previous study (Voors et al. 1978) only indicated an association in men between respiratory cancer and residence in wetland areas of Louisiana. Like prostate cancer, nonwhite population variable appeared to be an important deterministic factor for the stomach cancer of the combined sex and race. Nonwhite population tended to have low socioeconomic status, such as low educational level, low income, and low urban residence.

The stepwise regression analysis selected one variable [transportation (+)] among white males and two predictor variables [persons employed in health service (-) and total toxics releases (+)] among white females. This implied a potential health effect of exposure to toxic substances. Among nonwhite males, wetlands (+), persons

employed in mining (+), and chemical manufacturing (+) were the first, second, and third variable to enter the model, respectively.

In summary, this multiple regression analysis began with the hope of finding the causes and geographic distributions of cancer mortality rates and the environment for further study. The findings of this study were similar to those of the earlier studies. In the case of cancers of all sites combined, lung, and breast, strong, clear, and consistent causal relationship was found with one or two dominant environmental factors. For the cancers of colorecutm, prostate, and stomach, the causal relationship was weak, or significant, but inconsistent. Urban population had a positive effect on most cancer mortality rates. Lung cancer mortality rates were more closely related with occupational variables than any other cancer mortality rates. Breast cancer rates were positively affected by socio-economic variables. Colorectal cancer mortality rates had a positive relationship with Mississippi River as drinking water. Prostate cancer and stomach mortality rates were positively affected by nonwhite population. In particular, stomach cancer showed a positive relationship with wetland areas.

Even though this study used several risk factor data in Louisiana, the data were limited to the aggregate of parish scale. Accordingly, some statistical associations, although plausible, tended to be artifacts of the data, and this study might miss the true relationships between cancer mortality rates and environmental factors. Future study should be more concerned with this point. Nevertheless, this study provided additional information that was not visually evident from the maps and could help to test, narrow, and generate the hypotheses suggested by geographic patterns of various cancers and the environmental variables.

5.6 Summary

The purpose of this chapter was to examine the distribution of high cancer mortality rates and their relationships with environmental factors in Louisiana parishes. This chapter consists of five subsections and the results are summarized as follows:

1. The significance tests showed that cancer mortality rates for most parishes in South Louisiana were significantly higher than those of the U.S for cancers of all sites combined, lung, and stomach. Especially, most parishes in South Louisiana showed significantly higher cancer mortality rates for lung among males and stomach among nonwhite males than the U.S.

2. Factor analysis was performed on the 16 cancer mortality rates in Louisiana parishes from 1953 to 1977 and from 1978 to 1987. In both periods, the first factor represented cancers related to the digestive system and the second factor was considered as the lung cancer mortality rates for nonwhites. Even though the areas of high scores on major factors had been changed, they were dominantly in several parishes in South Louisiana or west regions of the lower Mississippi River.

3. The correlogram analysis of autocorrelation was undertaken to examine the spatial-temporal patterns of cancer mortality rates in Louisiana in 1953-1987, 1953-1977, and 1978-1987. For 35 years, cancers of all sites combined, lung, and stomach cancers exhibited the strongest degree of positive spatial autocorrelation, whereas cancers such as breast, colorectum, and prostate exhibited low spatial autocorrelation.

4. Before multiple regression, factor analysis was performed on 24 independent variables to find a more parsimonious set and to estimate the characteristics of the selected environmental variables. The environmental variables could be grouped into

physical environment, chemical pollution, socio-economic status, and demographic resource.

5. A multiple regression analysis was accomplished to find the relationship and causes of cancer mortality rates and environmental variables in Louisiana parishes for the 1980s. Most cancer morality rates had a significantly positive relationship with urban population and negative relationship with persons employed in health service and education service. There was a positive relationship between socio-economic activities and breast cancer rates and between drinking water and colorectal cancer rates. Lung cancer mortality rates were more closely related to occupational variables and exposure to toxic substances than any other types of cancer mortality rates. Prostate and stomach cancer mortality rates were significantly and positively related to nonwhite populations.

As time progresses, high cancer mortality rates have distributed more widely and randomly, but the mortality rates are still clustered significantly in the South. This chapter confirmed hypotheses that cancer mortality rates are higher in South Louisiana (than in the U.S.) and that there are spatial clusterings of cancer mortality rates of some major sites. Furthermore, this study provided additional information about the cancer mortality patterns associated with environmental variables.

CHAPTER 6

SMALLER THAN PARISH LEVEL

6.1 Scan Statistic Analysis of Lung Cancer in Louisiana from 1988 to 1993

Previous chapters and recent statistics (Miller et al. 1993) showed that the parishes (or counties) that have the highest age-adjusted lung cancer mortality rates among males are all located in the southeastern U.S. Cancer is one of the most serious health problems in Louisiana. In particular, approximately 85 % of cancer is attributed to lung cancer and the high rates of lung cancer mortality are prominent in the southern part of Louisiana.[6.1]

Based on the previous chapters, there has been spatial clustering of lung cancer mortality rates in Louisiana. A number of possible etiologic factors have also been suggested (i.e., lung cancer mortality rates were more closely related to occupational variables and exposures to toxic substances), some of which might explain the occurrence of clustering. However, the significance of such space, time and/or space-time clustering must be tested. There may be a problem in distinguishing between naturally occurring clusters because of chance alone, and those because of some underlying spatial risk factor. Also, it has been suggested that lung cancer may have different clusters at different scales. For these reasons, this chapter is to test and compare a large set of lung cancer deaths for the presence of geographical clusters in Louisiana parishes and census tracts from 1988 to 1993. As mentioned in Chapter 3, the section on Materials and Methods, the SaTScan software (which analyzes point data

[6.1] The elevated lung cancer rates in Louisiana men were first noted in the First National Cancers Survey of ten cities in 1937; the high rates has persisted in South Louisiana since 1950.

using the spatial, temporal, or space-time scan statistic) was used to determine if significant clusters of lung cancer exist.

6.1.1 Spatial Scan Statistic Analysis of Lung Cancer by Parish

Figure 6.1 shows age-adjusted lung cancer mortality rates in Louisiana during 1988-1993. The annual mortality rates of lung cancer were 61.5 deaths per 100,000, with a total of 16,162 deaths during the six years. It ranked the second in U.S. states (CDC Wonder 1999).

Scan statistic searched for clusters of cases without specifying their size or location and tested for their statistical significance while adjusting for the multiple testing inherent in such a procedure. For hypothesis testing, the SaTScan program generated a number of random replications of the data set under the null hypothesis (Monte Carlo procedure). The result is significant at the 0.05 level, for example, if the value of the test statistic from the real data set is among the 5% highest of all 1,000 values. Also, the log likelihood ratio is used to determine if an observed cluster is significant at the 0.01 and 0.05 levels. For example, an observed cluster (i.e., most likely cluster or secondary cluster) is significant if the log likelihood ratio level of the observed cluster is higher than the predetermined critical value.

Spatial, temporal, or space-time scan statistic was performed to consider the differences in cancer clusters in different areas and time. The SaTScan program tested for clusters with high, low, and high and low rates as well as for clusters by adjusting categorical covariates such as age, race, and sex. When more than one covariate was selected, each one was adjusted for as well as the interaction terms among them. The number of categorical covariates affected the geographical pattern of

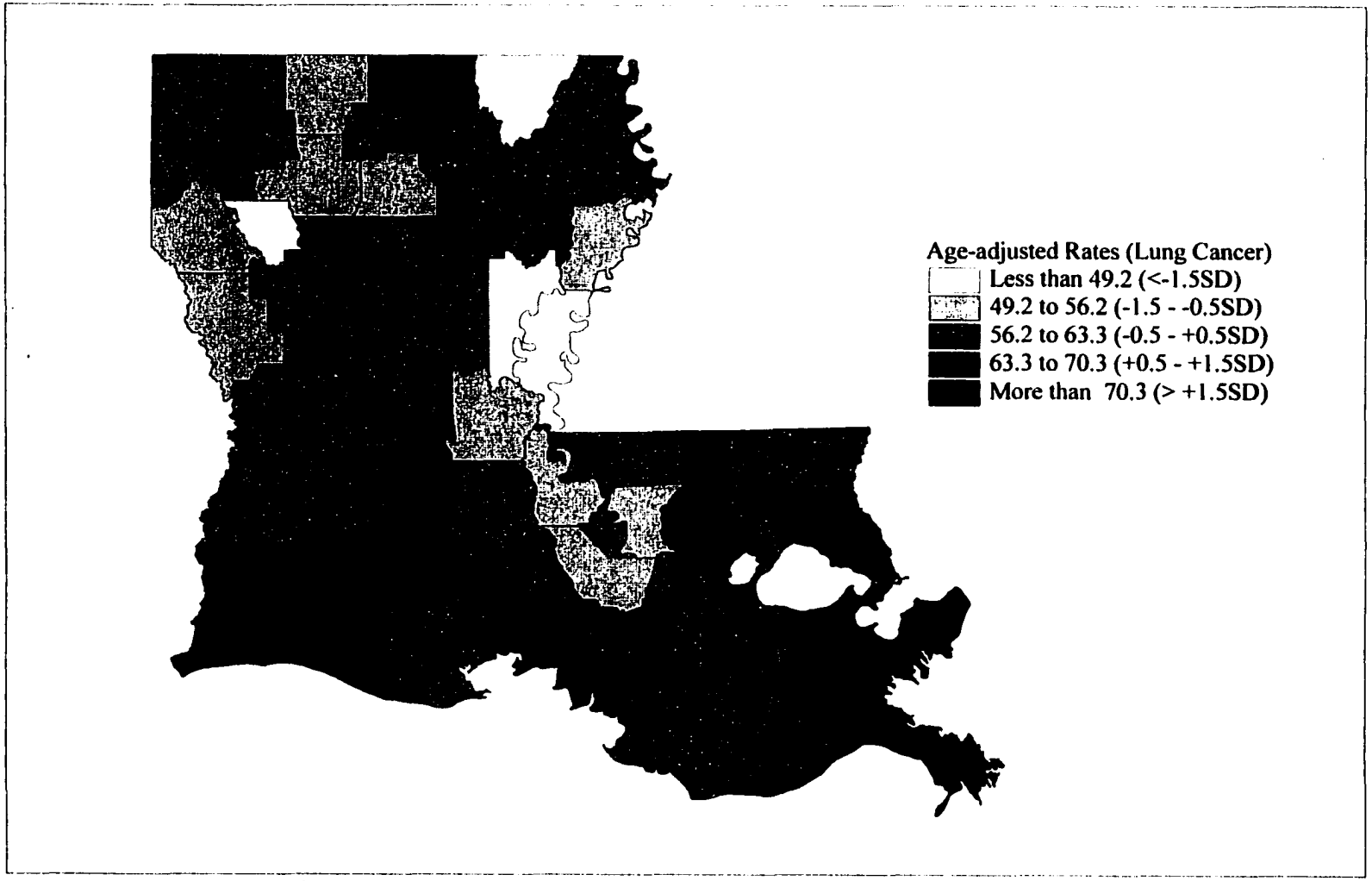

Figure 6.1 Lung Cancer Mortality Rates (All Races, Both Sexes): 1988-1993
(Source: Centers for Disease Control and Prevention, 1999)

mortality rates. This study detected the geographic clusters with high rates, based on the results of spatial scan statistic.

The results of spatial analysis for clusters with high lung cancer rates (by specifying age, race, and sex) are shown in Figure 6.2 and Table 6.1. The most likely cluster appeared in St. Bernard, Plaquemines, and Jefferson parishes. Its distribution represented geographically distinct areas which had higher lung cancer mortality rates than the rest of Louisiana. Table 6.1 shows that the most likely cluster was a statistically significant cluster at the level $p = 0.001$.

In addition to the most likely cluster, the method also identified secondary clusters in the data set, and ordered them according to their likelihood ratio. There were two secondary clusters that did not overlap the most likely cluster. The secondary clusters were centered in eight parishes of southwestern Louisiana (Vermilion, Lafayette, Acadia, Iberia, Jefferson Davis, St. Martin, St. Mary, and Cameron) and in a small area of Livingston Parish. However, as can be seen in Table 6.1, neither of these was statistically significant (with p-values of 0.191 and 0.9770). That lack of significance could be because the result has occurred truly by chance or because the increased risk and the power of the test are too low to detect it (Kulldorff et al. 1997). Usually, the SaTScan program does not report clusters of this type since most of them provide little additional information, but their existence means that while it is possible to pinpoint the general location of a cluster, its exact boundaries must remain uncertain (SaTScan V.2.1). Furthermore, Vermilion (58.4) Parish within the secondary cluster showed lower mortality rates than the average rates (61.5) of Louisiana.

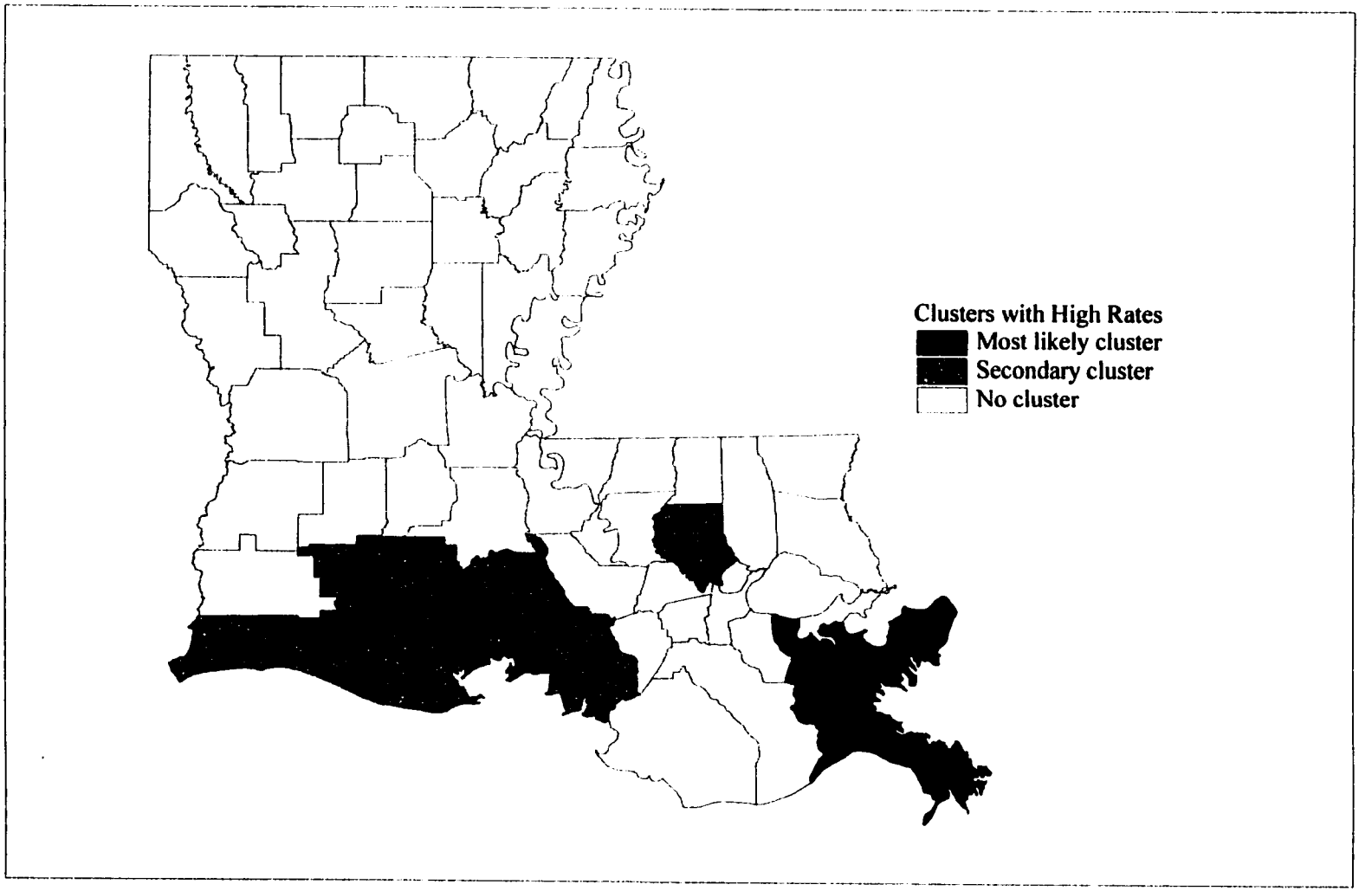

Figure 6.2 Spatial Clusters of Lung Cancer by Parish, Louisiana (1988-1993)

Table 6.1 Spatial Clusters of Lung Cancer in Louisiana Parishes, 1988-1993, Using the Spatial Scan Statistic

	SA (H) Type	Location	Cases	Expected	Log likelihood ratio	*P*-value
Total cases	M	Plaquemines, St. Bernard, Jefferson	2179	1936.13	16.72	0.001
	S	(a)Vermilion, Lafayette, Acadia, Iberia, Jefferson Davis, St. Martin, St. Mary, Cameron	1845	1720.31	4.96	0.191
		(b) Livingston	248	219.20	1.84	0.977
White males	M	St. John Baptist, St. James, St. Charles, Ascension, Livingston, Assumption, Jefferson, Lafourche	1573	1464.53	4.87	0.205
	S	(a) Allen, Evangeline, Lafayette, Beauregard, Vernon	528	476.90	2.83	0.778
		(b) Ouachita, Richland, Jackson, Caldwell, Union	464	425.82	1.77	0.985
White females	M	Plaquemines, St. Bernard, Jefferson	702	595.83	10.62	0.001
	S	(a) Cameron, Calcasieu, Jefferson Davis	265	220.43	4.49	0.246
		(b) Caddo, Bossier	379	337.61	2.67	0.821
		(c) St. Mary, Iberia	127	107.35	1.75	0.991
Nonwhite males	M	(a) Vermilion, Lafayette, Acadia, Iberia, Jefferson Davis, St. Martin, St. Mary, Cameron, St. Landry, Evangeline	454	370.70	9.99	0.005
	S	(a) Jefferson, St. Charles, Orleans, Lafourche, St. Bernard, Plaquemines, St. John Baptist, St. James, St. Tammany, Terrebonne, Ascension, Assumption, Livingston	1252	1157.19	16.07	0.076
		(b) Union	28	20.83	1.12	0.999

Spatial analysis (SA) for clusters with high rates (H), specifying age, race, and sex
Most likely cluster (M), Secondary cluster (S)

(Table 6.1 continued)

Nonwhite females	M	Lafourche, St. Charles, Jefferson, Terrebonne, St. John Baptist, St. James, Assumption, Orleans, Plaquemines, Ascension, St. Bernard, St. Mary, Livingston, St. Tammany, Iberville, Tangipahoa	601	509.38	13.69	0.001
	S	Cameron, Calcasieu, Jefferson Davis, Vermilion, Beauregard, Bienville, Lafayette	131	96.34	6.12	0.060

In addition to comparing total lung cancer deaths, race and sex-specific lung cancer deaths were examined.

White Males: Lung cancer mortality rates among white males in Louisiana, 91.8 deaths per 100,000, were the fifth highest in nation (Figure 6.3 and Table 6.1). Although one most likely cluster and two secondary clusters were detected, they were not significant, with *p*-values of 0.205, 0.778, and 0.985, respectively. The parishes in the most likely cluster (St. John Baptist, St. James, St. Charles, Ascension, Livingston, Assumption, Jefferson, and Lafourche) and secondary clusters (Allen, Evangeline, Lafayette, Beauregard, Vernon, Ouachita, Richland, Jackson, Caldwell, and Union) had relatively higher cancer mortality rates, compared with those of other parishes (except for Lafourche (90.0) and Jefferson Davis (87.6)).

White Females: Louisiana's lung cancer mortality rates among white females, 36.0 deaths per 100,000 population, ranked the ninth in the U.S. rates. The results of scan statistic showed one most likely cluster and three secondary clusters (Figure 6.4). Like the cluster of all lung cancer deaths, the cluster of lung cancer among white females centered in Plaquemines, St. Bernard, and Jefferson parishes in southeastern Louisiana. The most likely cluster was statistically significant ($p = 0.001$), whereas the secondary clusters were not significant. However, the existence of the secondary clusters implied the general location of a cluster, even though its exact boundaries were not clear. Clusterings of white females were smaller and more random, compared with those of other sex-races.

Nonwhite Males: As shown in Figure 6.5, this scan statistic indicated one most likely cluster and two secondary clusters. In general, there was a statistically significant

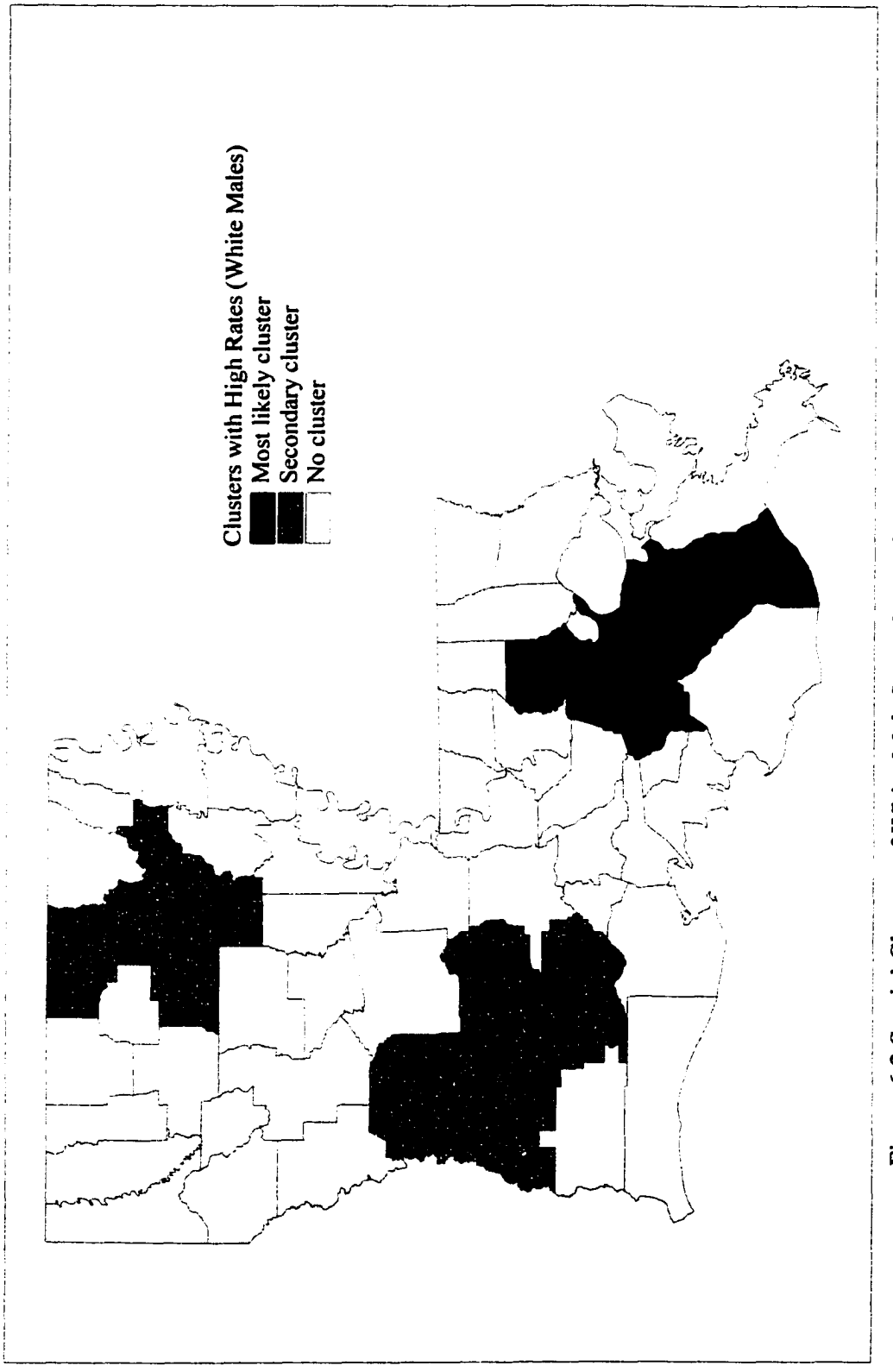

Figure 6.3 Spatial Clusters of White Male Lung Cancer by Parish, Louisiana (1988-1993)

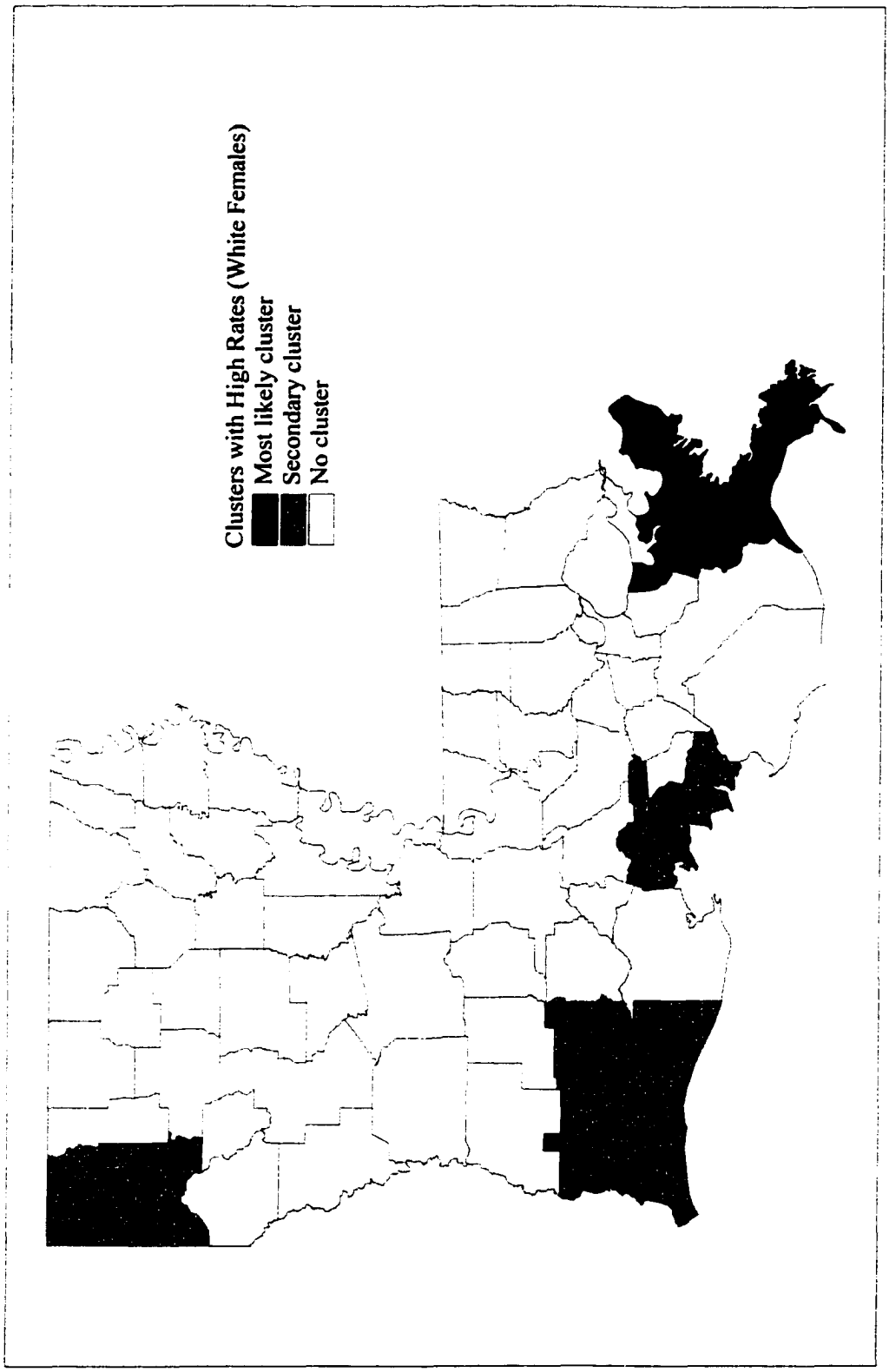

Figure 6.4 Spatial Clusters of White Female Lung Cancer by Parish, Louisiana (1988-1993)

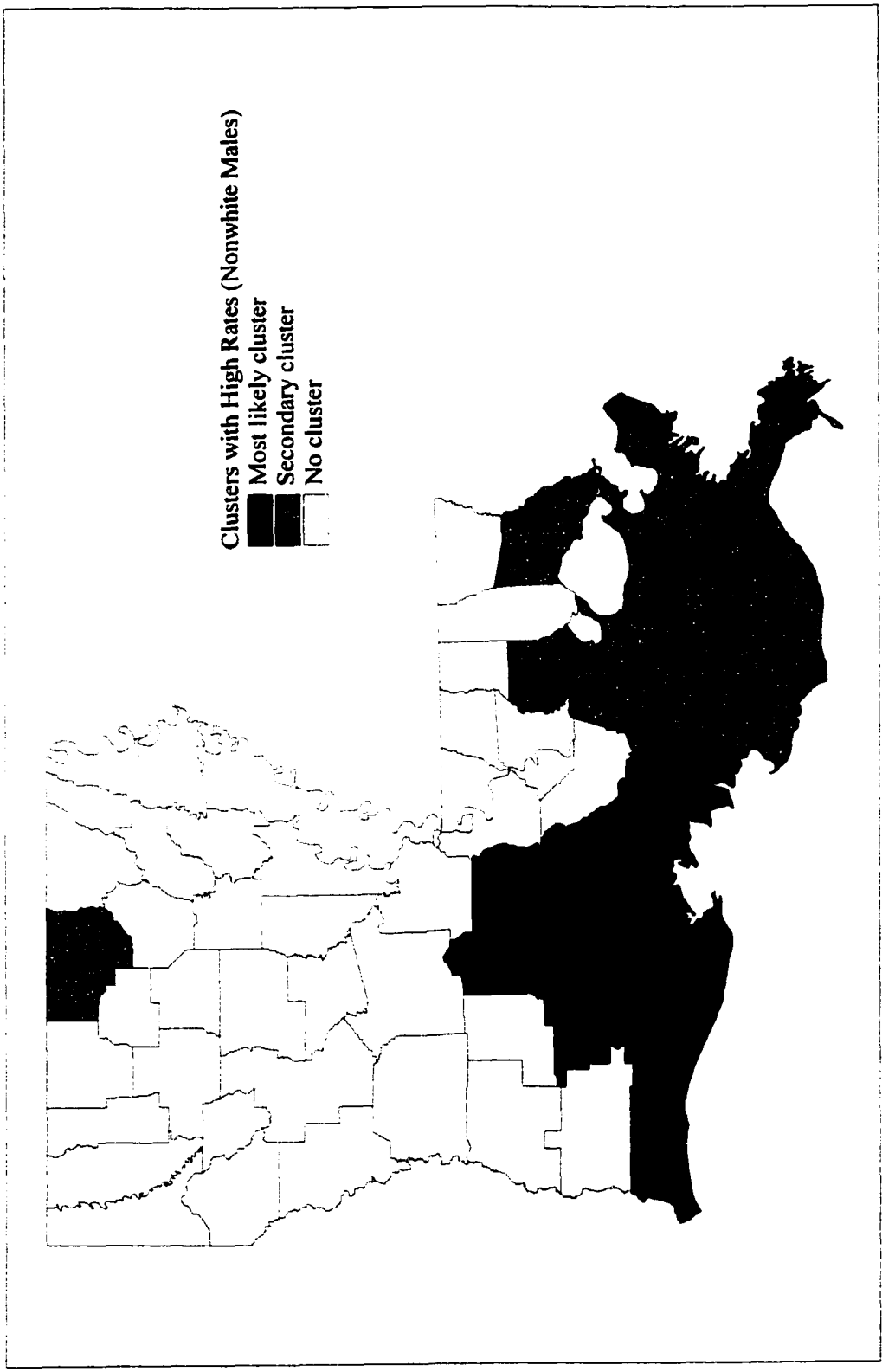

Figure 6.5 Spatial Clusters of Nonwhite Male Lung Cancer by Parish, Louisiana (1988-1993)

($p = 0.005$) and geographically broad cluster of lung cancer deaths among nonwhite males in Vermilion, Lafayette, Acadia, Iberia, Jefferson Davis, St. Martin, St. Mary, Cameron, St. Landry, and Evangeline parishes. Unlike the most likely cluster of sex-race combined or white lung cancer deaths, the distribution of the most likely cluster for nonwhite males appeared in southwestern Louisiana, with Acadia Parish as the center.

The two secondary clusters centered in broad areas (13 parishes) in southeastern Louisiana and one small parish (Union) in northern Louisiana. Neither of these was statistically significant (with p-values of 0.076 and 0.999). According to previous studies and statistics, the distributions of high cancer mortality rates among nonwhite males were widely concentrated in southern Louisiana.

Nonwhite Females: As depicted in Figure 6.6, the most likely cluster and the secondary cluster were rather broadly presented in southern Louisiana. Their distributions tended to be opposite to those of nonwhite males. In other words, the most likely cluster was centered in southeastern Louisiana whereas the secondary cluster was generated in southwestern Louisiana. The most likely cluster was statistically significant ($p = 0.001$); on the other hand, the secondary cluster was not significant ($p = 0.060$). The areas of most likely cluster included Lafourche, St. Charles, Jefferson, Terrebonne, St. John Baptist, St. James, Assumption, Orleans, Plaquemines, Ascension, St. Bernard, St. Mary, Livingston, St. Tammany, Iberville, and Tangipahoa parishes. For nonwhite females, the distribution of high death rates was generally random in 1950s. It has shown clustering of high death rates in the South since 1960, and the high death rates were particularly clustered in the southeastern area from 1988 to 1993.

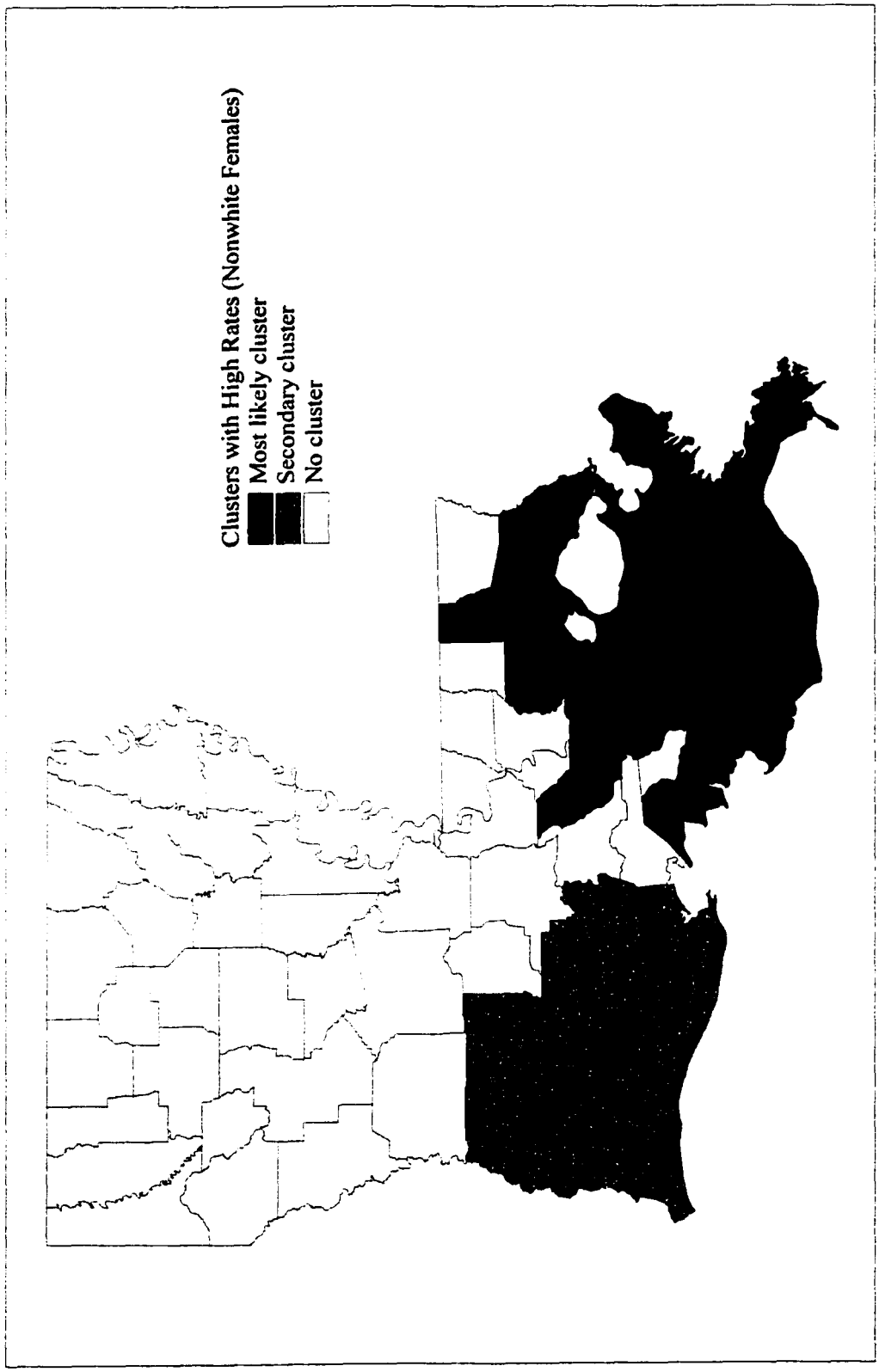

Figure 6.6 Spatial Clusters of Nonwhite Female Lung Cancer by Parish, Louisiana (1988-1993)

Although this section did not list the clusters of males, females, whites, and nonwhites separately, lung cancer among males and nonwhites had broader clusters with high rates than those among females and whites. This analysis showed that in general, the clusters of sex and race-specific lung cancer deaths (Table 6.1) were found in southeastern Louisiana, except for the cluster of lung cancer death among nonwhite males.

The significant cluster for all lung cancer deaths by parish in Louisiana could be characterized in the following. Given the environmental variables used in previous multiple regression analysis, it was suggested that the parishes (St. Bernard, Plaquemines, and Jefferson) in the most likely cluster of lung cancer had the largest percentages of persons employed in transportation in Louisiana (Figure 6.7). Most people in these three parishes lived in urban areas and their parishes showed high urban population rates, 96.4, 68.7, and 93.6, respectively (Figure 6.8). This cluster area showed the lowest person below poverty level (Figure 6.9) and highest per capita income. Except for Plaqumines, St. Bernard and Jefferson had a high population density. All three parishes used Mississippi River as drinking water. The parishes released high amounts of total TRI (Figure 2.7) as well as carcinogenic TRI (Figure 6.10). In particular, they had high air TRI emissions. A significant source of these pollutants has been attributed to a heavy concentration of industry along the Mississippi River, and discharges from traffic. Furthermore, more than half of the parishes were uninhabitable wetlands (Figure 6.11). Wetlands are important to wildlife and for maintaining water quality. Cancer might be a potential health effect of exposure to toxic substances in wetlands (Voors et al. 1978). In general, the significant most likely

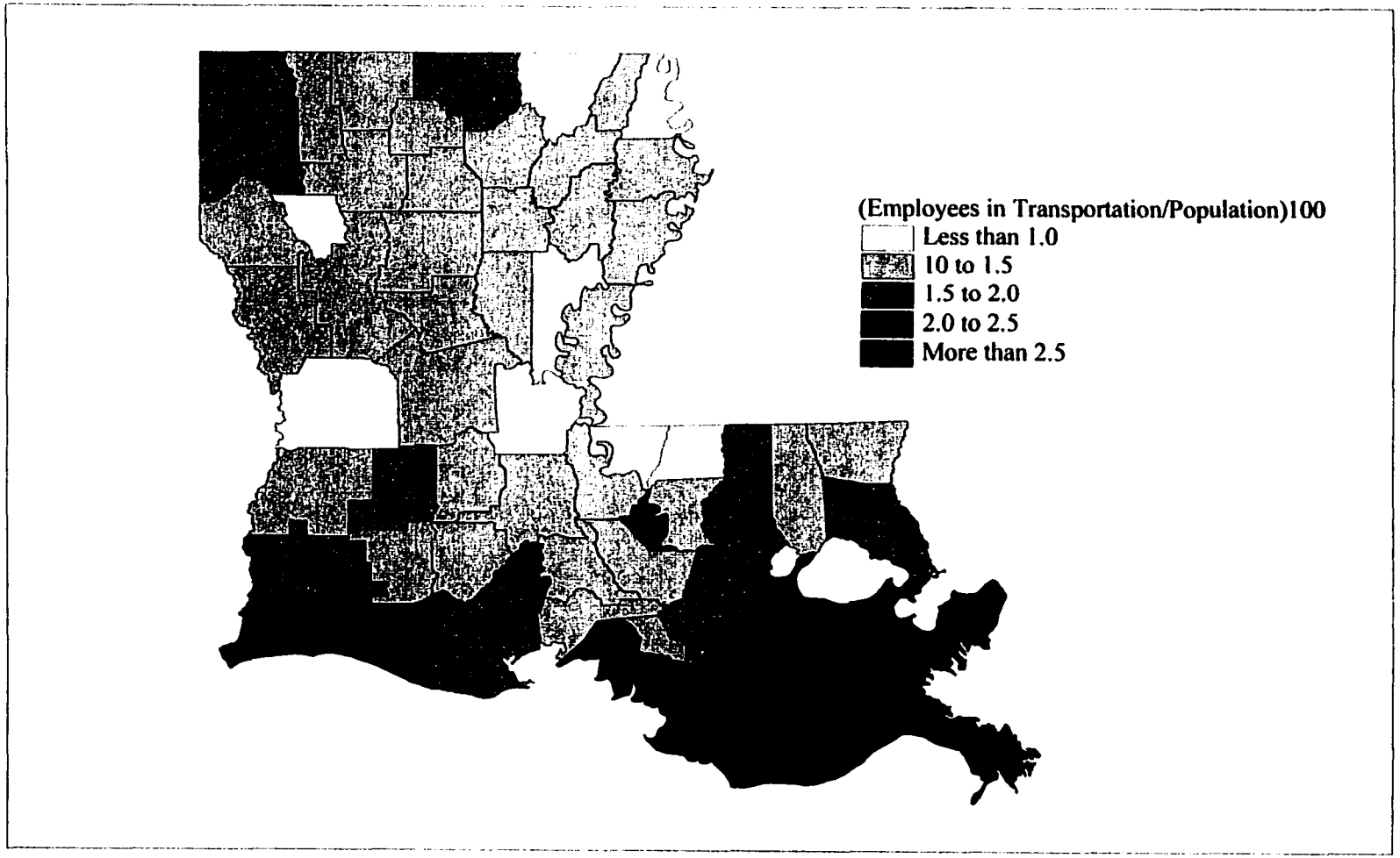

Figure 6.7 Employment in Transportation, Louisiana (1980s)
(Source: Regional Economic Information Systems and County and Business Pattern, 1980-1989)

Figure 6.8 Urban Population in Louisiana (1980s)
(Source: Louisiana FactBook, 1993)

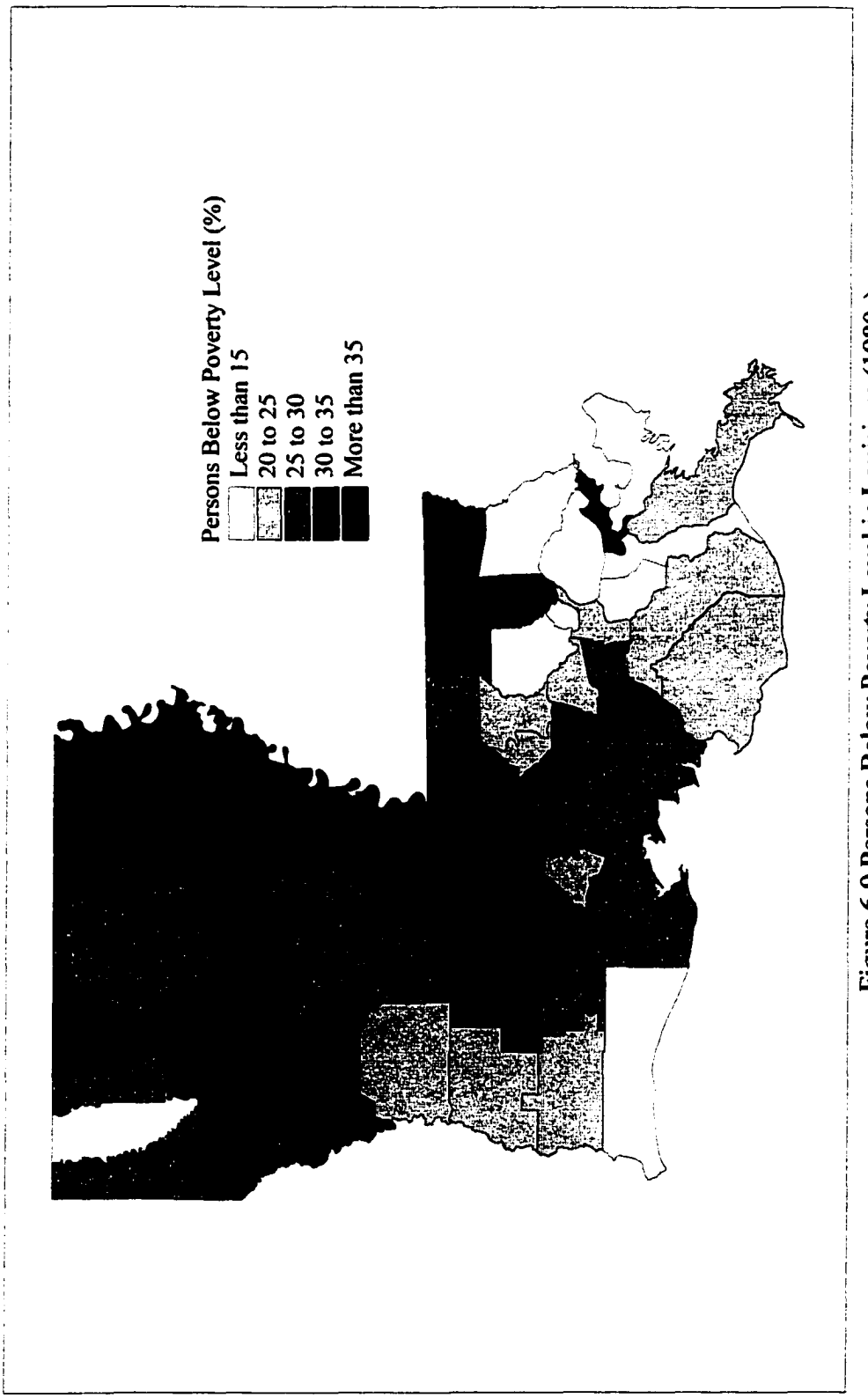

Figure 6.9 Persons Below Poverty Level in Louisiana (1980s)
(Source: U.S. Census Bureau, 1984 and 1884)

Figure 6.10 Total Carcinogenic TRI Releases, Louisiana (1987-1989)
(Source: Louisiana Department of Environmental Quality, 1993)

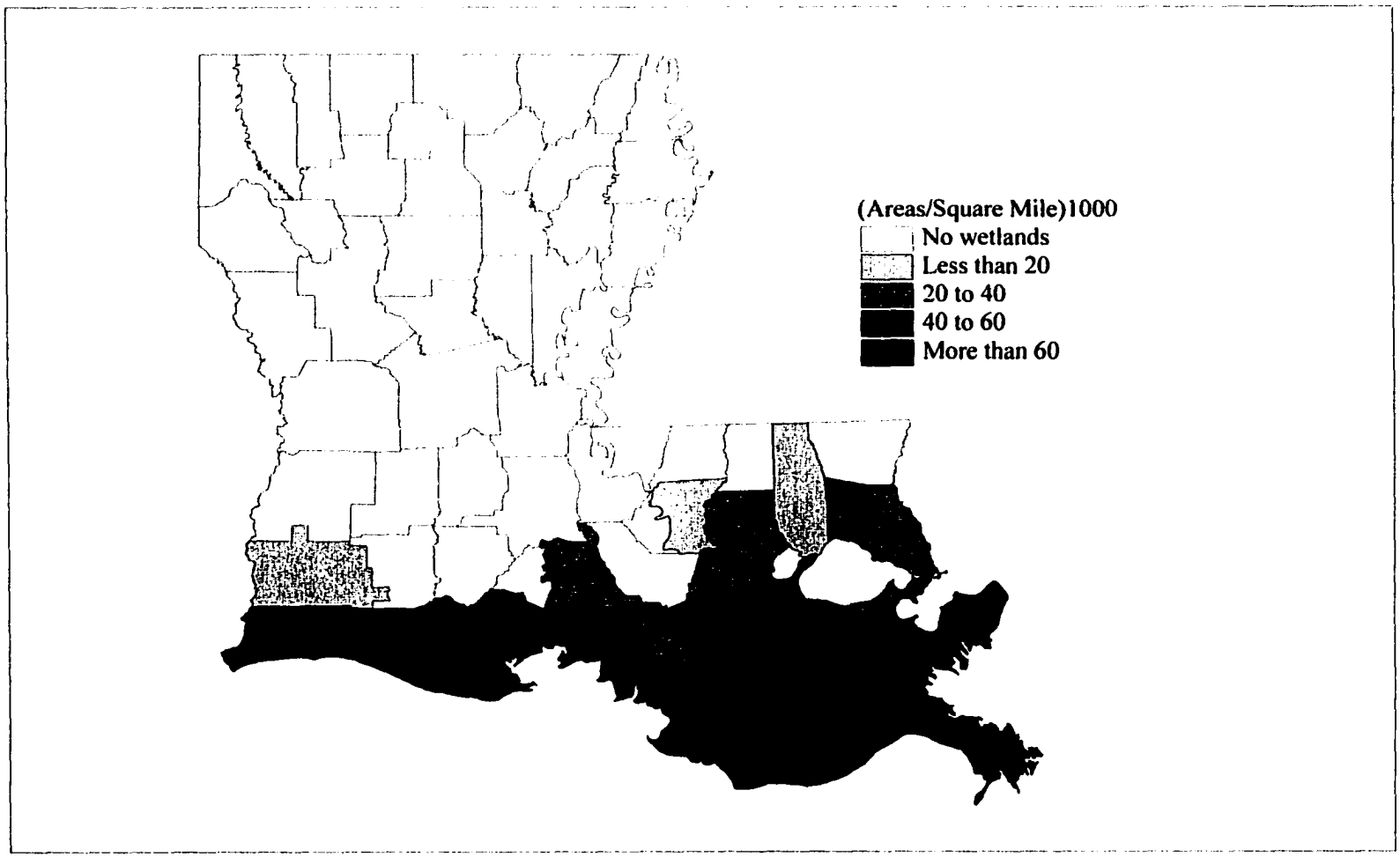

Figure 6.11 Wetlands in Louisiana (1980s)
(Source: Calculated by author from the Classified Habitat Data of Louisiana Department of Natural Resource, 1978 and 1988)

cluster of lung cancer for the period 1988-1993 could be characterized by urbanization and industrialization in wetlands.

Even though the most likely cluster among white males was not significant, it could be shown that the cluster includes most Mississippi River industrial corridor parishes. Parishes in significant most likely cluster of nonwhite males showed relatively low education status (except for Lafayette) (Figure 6.12) and low rates of urban population. Most parishes were included in the Acadiana region (except for Cameron and Jefferson Davis) and had dominantly high percentages of persons employed in mining (Figure 6.13). Acadiana was settled by Catholics who fled persecution in Nova Scotia. Much of the population of Acadiana region is of French/Cajun descent. This cluster could be characterized by industrialization in the Acadiana region. The cluster of lung among white females was the same as that of total lung cancer and that among nonwhite females was similar to that of total lung cancer.

In addition to investigating clustering (by age, sex, and race) on sex and race-specific lung cancer deaths, this study adjusted for any number of categorical covariates by specifying age, sex, and race, separately or together (Table 6.2). The SaTScan program searched for clusters above and beyond that which is expected due to the covariates. As mentioned before, when covariates were specified, each one was adjusted for as well as the interaction terms between them.

By specifying age and race as categorical covariates, this method found the same most likely cluster, as in clusters adjusted before by age, sex, and race. The cluster was significant at the level $p = 0.001$. There were two secondary clusters. The secondary cluster found in the southwestern Louisiana was broader than the cluster

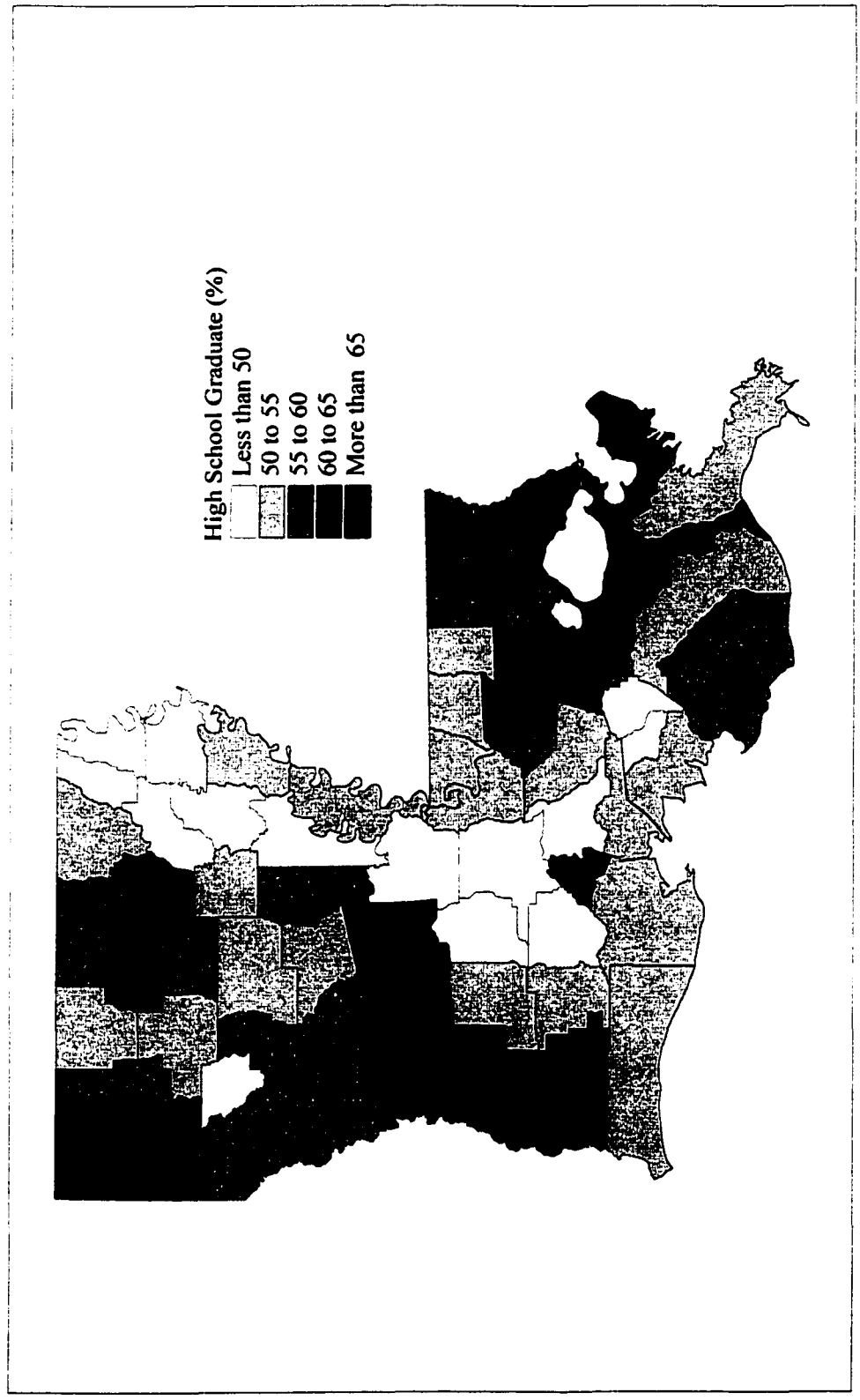

Figure 6.12 Education Status, Louisiana (1980s)
(Source: U.S Census Bureau, 1984 and 1994)

Figure 6.13 Employment in Mining, Louisiana (1980s)
(Source: Regional Economic Information Systems and County and Business Pattern, 1980-1989)

Table 6.2 Spatial Clusters of Lung Cancer in Louisiana Parishes (by Covariate), 1988-1993, Using the Spatial Scan Statistic

Covariate	SA (H) Type	Location	Cases	Expected	Log likelihood ratio	*P*-value
Age groups Race, & Sex	M	Plaquemines, St. Bernard, Jefferson	2179	1936.13	16.72	0.001
	S	(a)Vermilion, Lafayette, Acadia, Iberia, Jefferson Davis, St. Martin, St. Mary, Cameron	1845	1720.31	4.96	0.191
		(b) Livingston	248	219.20	1.84	0.977
Age groups & race	M	Plaquemines, St. Bernard, Jefferson	2179	1942.64	15.81	0.001
	S	(a) Vermilion, Lafayette, Acadia, Iberia, Jefferson Davis, St. Martin, St. Mary, Cameron, St. Landry, Evangeline, Allen, Iberville, Calcasieu	3232	3050.36	5.57	0.043
		(b) Livingston	248	214.26	2.56	0.860
Age groups & sex	M	Plaquemines, St. Bernard, Jefferson, Orleans, Lafourche	4597	4337.65	10.48	0.004
	S	(a) Vermilion, Lafayette, Acadia, Iberia, Jefferson Davis, St. Martin, St. Mary, Cameron, St. Landry, Evangeline, Allen, Iberville, Calcasieu	3235	3079.37	4.80	0.222
Race & sex	M	St. Bernard, Orleans	2447	2009.87	51.27	0.001
	S	(a) Claiborne, Webster, Lincoln, Union Bienville, Bossier, Jackson, Caddo, Red River, Ouachita, Winn, De Soto, Morehouse, Natchitoches, Richland, Grant, West Carroll, Franklin, Sabine, La Salle, East Carroll	3625	3339.00	15.14	0.001

Spatial analysis (SA) for clusters with high rates (H), Most likely cluster (M), Secondary cluster (S)

(Table 6.2 continued)

	S	(b) Jefferson Davis, Acadia, Allen, Calcasieu, Cameron, Evangeline	1373	1254.52	5.96	0.084
		(c) St. Mary, Iberia, Assumption	609	565.70	1.67	0.989
Age groups	M	Lafourche, St. Charles, Jefferson, Terrebonne, St. John Baptist, St. James, Assumption, Orleans, Plaquemines, Ascension, St. Bernard, St. Mary, Livingston, St. Tammany	6428	6177.28	8.21	0.009
	S	Cameron, Calcasieu, Jefferson Davis, Vermilion, Beauregard, Acadia, Allen, Lafayette, Evangeline, Vernon, St. Landry, Iberia	3009	2840.32	5.90	0.070
Race	M	St. Bernard, Orleans	2447	2027.16	47.08	0.001
	S	(a) Caddo, Bossier, Webster, De Soto, Red River, Bienville, Claibone, Lincoln, Sabine, Natchitoches, Jackson, Winn, Union, Ouachita, Grant, Caldwell, Vernon, La Salle, Morehouse, Rapides, Richland, Franklin, Catohoula, Beauregard, West Carroll, Allen, Avoyelles, Concordia, Tensas, East Carroll, Madison, Evangeline, Calcasieu, Jefferson Davis, St. Landry, Acadia, Pointe Coupee	6492	6168.56	13.65	0.001
Sex	M	St. Bernard, Orleans	2450	2141.48	24.66	0.001
	S	(a) De Soto, Red River, Sabine, Caddo, Bossier, Natchitoches, Bienville, Webster, Winn, Claiborne, Jackson, Lincoln, Vernon, Grant, Caldwell, La	6533	6127.41	21.48	0.001

(Table 6.2 continued)

		Salle, Rapides, Union, Ouachita, Allen, Beauregard, Catahoula, Richland, Evangeline, Franklin, Avoyelles, Morehouse, Calcasieu, Concordia, Jefferson Davis, Tensas, West Carroll, St. Landry, Acadia, Madison, East Carroll, Cameron, Pointe Coupee				
Only cases	M	St. Bernard, Orleans	2460	2147.46	25.22	0.001
		Caddo, Bossier, De Soto, Webster, Red River, Bienville, Claiborne, Sabin, Natchitoches, Lincoln, Jackson, Win, Union, Vernon, Grant, Ouachita, Caldwell, Rapides, La Salle, Beauregard, Morehouse, Richland, Allen, Catahoula, Franklin, Evangeline, Calcasieu, Avoyelles, West Carroll, Concordia, Jefferson Davis, Tensas, Madison, East Carroll, Acadia, St, Landry, Cameron, Point Coupee	6554	6167.46	19.44	0.001
		St. Mary, Iberia, Assumption	611	567.47	1.69	0.988

adjusted by age, sex, and race. Unlike the cluster by age, sex, and race, this cluster was rather significant ($p = 0.043$). Another cluster (in Livingston Parish) of the two secondary clusters was the same and was not significant, compared with clusters adjusted by age, sex, and race.

Specifying age and sex as categorical covariates generated a significant most likely cluster ($p = 0.004$) and an insignificant secondary cluster. The most likely cluster included Plaquemines, St. Bernard, Jefferson, Orleans, and Lafourche parishes.

When specifying race and sex, the results of spatial scan statistic showed one most likely cluster and three secondary clusters. The most likely cluster centered at St. Bernard and Orleans, which were smaller in size but was statistically significant ($p = 0.001$). The three secondary clusters were shown in northern, southeastern, and mid-southern Louisiana. Only, the secondary cluster in the northern Louisiana was statistically significant ($p = 0.001$).

In case of age groups alone, the distributions of the most likely cluster and the secondary cluster were more widely distributed in southeastern and southwestern Louisiana. Only the most likely cluster was statistically significant ($p = 0.009$) and the secondary cluster was not significant.

When specifying for race, sex, and case, separately, the results of each confounding variable were very similar among covariates. According to each covariate, the only most likely cluster was statistically significant in St. Bernard and Orleans ($p = 0.001$). All the secondary clusters were not significant and their distribution consistently appeared in most of northern Louisiana.

In general, the results of analyses indicated that adjusting for age group or incorporating age group among other covariates played an important role on the clustering patterns of lung cancer mortality rates. That is to say, if the age distribution is not the same in different areas, there is geographical clustering of lung cancer simply due to the age covariate. Existence for the older age groups (at ages 50 and over) not only increased cancer rates but also affected the geographic clustering of the cancer, while cancer rates for the younger age groups remained relatively stable and similar among parishes.

6.1.2 Spatial Scan Statistic Analysis of Lung Cancer by Census Tract

This section further examined the spatial clusters of high lung cancer deaths in Louisiana by using the census tract level data. Using the same procedures previously described, 22 clusters (the one most likely cluster and 21 secondary clusters) emerged. They are shown in Figure 6.14 and Table 6.3. This analysis was accomplished by using lung cancer cases and population only, because variables on age, sex, and race for each case are not available.

The most likely cluster was significant at the level $p = 0.001$ and its distribution was distinct and broad in southeastern Louisiana. The cluster included a total of 254 census tracts: 79 tracts in Jefferson, 170 tracts in Orleans, and 5 tracts in St. Bernard. Close inspection of Figure 6.15 reveals that these census tracts were distributed along the Mississippi River in the New Orleans region (Jefferson, Orleans, and St. Bernard parishes).

This study did not have all the environmental data to efficiently detect and examine clusters at the tract level, except for population information, TRI data, and the

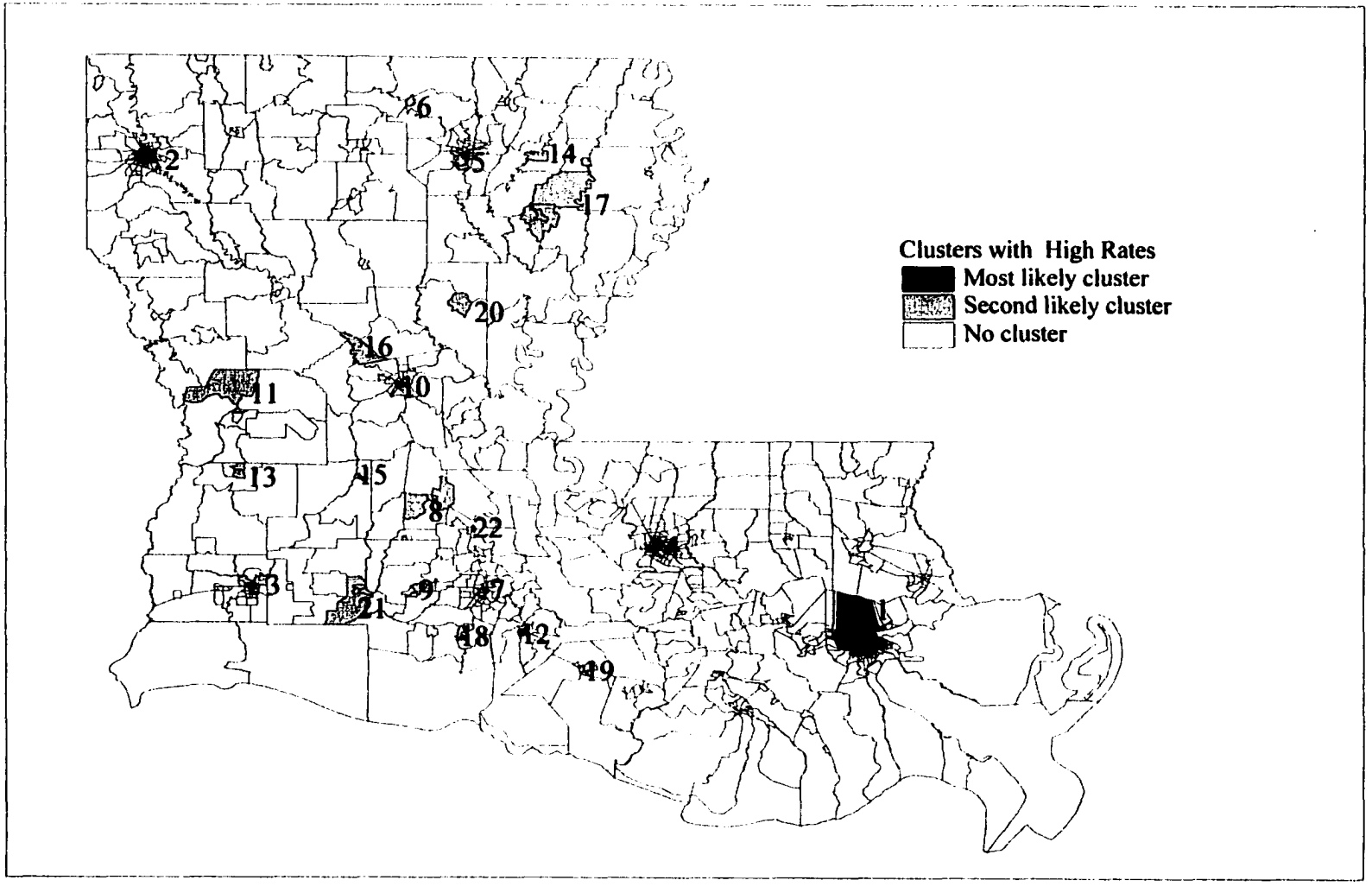

Figure 6.14 Spatial Clusters of Lung Cancer by Census Tract, Louisiana (1988-1993) (Mihye Bark, 1999)

Table 6.3 Spatial Clusters of Lung Cancer in Louisiana Census Tracts, 1988-1993, Using the Spatial Scan Statistic

	SA (H) Type	Location	Cases	Expected	Log likelihood ratio	*P*-value
1	M	22051-Jefferson (79), 22071-Orleans (170), 22087-St. Bernard (5)	3142	2191.4	222.1	0.001*
2	S(2)	22015-Bossier (6), 22017-Caddo (35)	567	377.0	42.8	0.001*
3	S(3)	22019-Calcasieu (12)	226	128.5	30.5	0.001*
4	S(4)	22033-East Baton Rouge (39)	593	433.1	27.4	0.001*
5	S(5)	22073-Ouachita (13)	177	970	26.7	0.001*
6	S(6)	22111-Union (1)	40	11.1	22.4	0.001*
7	S(7)	22055-Lafayette (7)	159	93.9	18.7	0.001*
8	S(8)	22039-Evangeline (4)	107	58.7	16.0	0.001*
9	S(9)	22001-Acadia (2)	70	31.9	12.4	0.004*
10	S(10)	22079- Rapides (6)	97	56.3	12.2	0.006*
11	S(11)	22115-Vernon (2)	56	28.89	9.98	0.042*
12	S(12)	22045-Iberia (5)				0.05*
13	S(13)	22011-beauregard (2)				0.141
14	S(14)	22083-Richland (1)				0.487
15	S(15)	22003-Allen (1)				0.545
16	S(16)	22043-Grant (1)				0.552
17	S(17)	22041-Frankline (3), 22083-Richland (1)				0.581
18	S(18)	22113-Vermilion (3)				0.642
19	S(19)	22101-St. Mary (3)				0.866
20	S(20)	22059- La Salle (1)				0.936
21	S(21)	22053-Jefferson Davis(4)				0.960
22	S(22)	22097-St. Landry (1)				0.982

Spatial analysis (SA) for clusters with high rates (H), Most likely cluster (M), Secondary cluster (S)
The numbers in the parenthesis are tract numbers, *: Significant at 0.05
See Appendix U.

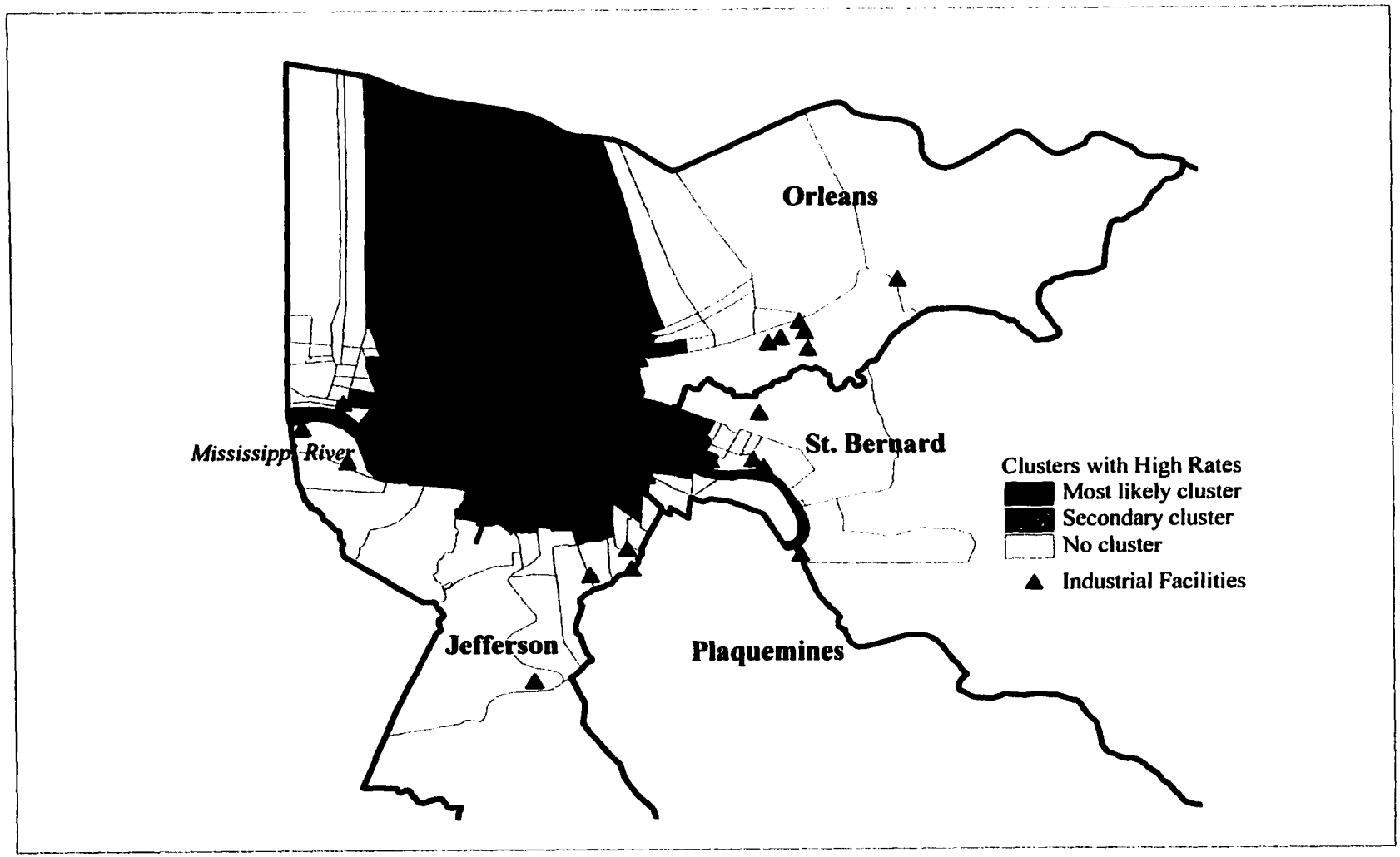

Figure 6.15 Spatial Clusters of Lung Cancer by Census Tract, New Orleans Region (1988-1993) (Miihye Bark, 1999)
(Industrial Facilities Source: Louisiana Department of Environment Quality, 1990)

distribution of wetlands. In general, these census tracts had a high population density, high TRI emissions from petrochemical manufacturing facilities and transportation, and wide areas of wetlands nearby. Residensts in these tracts used the Mississippi River as drinking water.

The secondary clusters were widely distributed in Louisiana. Eleven clusters were statistically significant at the level $p = 0.05$ or smaller and 10 clusters were not. The first cluster in the secondary clusters, called Cluster (2), is presented in Figure 6.14, as number 2 (2 in Figure 6.14). The cluster centered at 6 tracts in Bossier Parish and 35 tracts in Caddo Parish. These 41 tracts were a significant subcluster, with $p = 0.001$. Cluster (3) included 12 tracts in Calcasieu was distributed in southwestern Louisiana. Cluster (4) was located in 39 tracts of East Baton Rouge and Cluster (5) was found in 13 tracts of Ouachita. Clusters 6, 7, and 8 appeared in Union (1 tract), Lafayette (7 tracts), and Evangeline (4 tracts). Clusters 9, 10, and 11 included 2 tracts in Acadia, 6 tracts in Rapides, and 2 tracts in Vernon.

Previous studies reported that the significance levels of the secondary clusters were sometimes overestimated with simulated conservative p-values (i.e., p-values are smaller than their true values) (Kulldorff 1997). Therefore, careful interpretation is needed for small subclusters of the secondary cluster, even though it is statistically and spatially significant. Ten clusters were not explained here because they were not statistically significant.

Clusters (2), (3), (4) (in Bossier and Caddo, Calcasieu, and Baton Rouge) had similar characteristics to those in the most likely cluster. In Figure 6.16, Cluster (4) in East Baton Rouge Parish is shown to be around a petrochemical manufacturing corridor

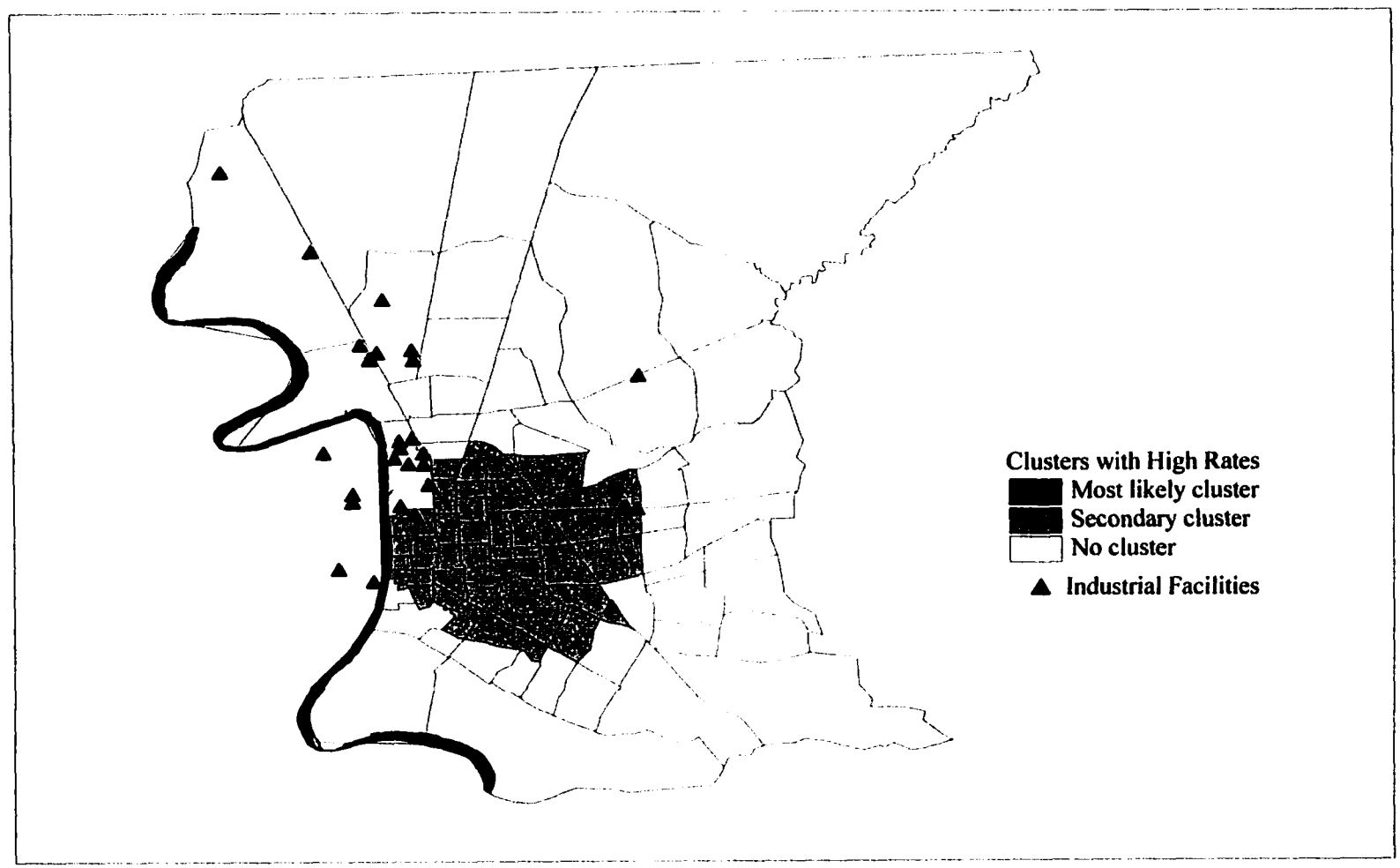

Figure 6.16 Spatial Clusters of Lung Cancer by Census Tract, Baton Rouge (1988-1993) (Miihye Bark, 1999)
(Industrial Facilities Source: Louisiana Department of Environment Quality, 1990)

along the Mississippi River. East Baton Rouge Parish, where the state capital is located, had a high per capita income, a big urban population (Figure 6.8), low person below poverty level (Figure 6.9), and the highest percentage of high education status (Figure 6.12). This parish released high amounts of TRI (Figure 2.7) as well as carcinogenic TRI (Figure 6.10).

6.1.3 Comparison of Spatial Clusters of Lung Cancer at the Different Scales

As mentioned earlier, the spatial scan statistics based on data at different spatial scales may lead to different results. The purpose of this section is to compare and clarify a large set of lung cancer deaths for the presence of geographical clusters at both scales (Louisiana parishes and census tracts) from 1988 to 1993. This study conditioned the analysis was based on the total number of cases observed only, because of unavailable data. The expected number of cases in each area (at two scales) under the null-hypothesis was calculated without specifying age, race, or sex as categorical covariates. To get correct significance levels that are neither conservative nor liberal, the number of Monte Carlo replications used must be adjusted for different data sets (small and large). More replications give higher power to the test results.

When comparing the geographic clusters of lung cancer deaths between the two scales, the parish level generated one most likely cluster and two secondary clusters (Figure 6.17), whereas the census tract level developed one most likely cluster and twenty-one secondary clusters (Figure 6.14).

At the parish level, the most likely cluster was shown in St. Bernard and Orleans of southeastern Louisiana. The cluster was statistically significant ($p = 0.001$). Areas of the two secondary clusters appeared broadly in most parishes of northern Louisiana

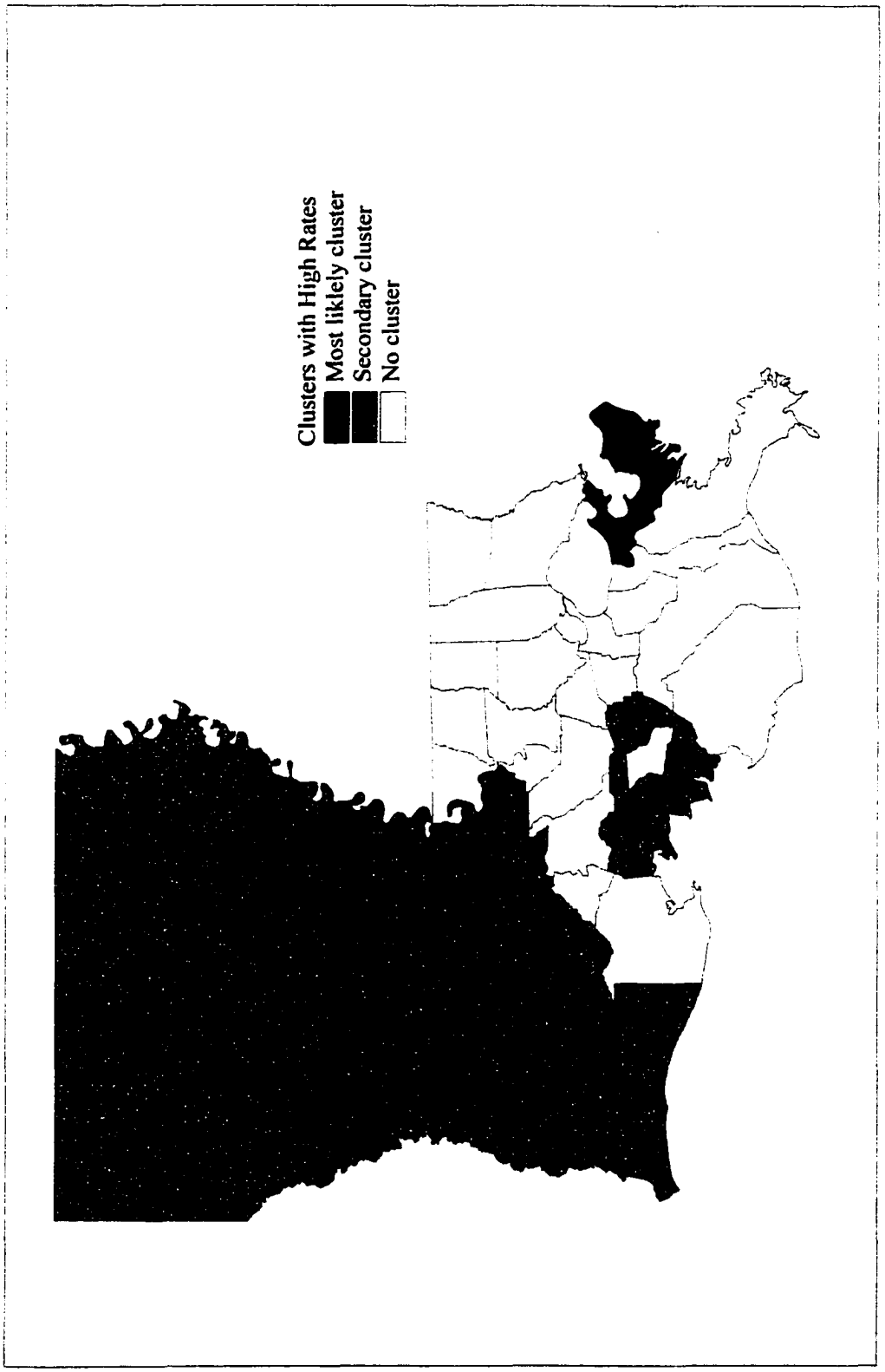

Figure 6.17 Spatial Clusters of Lung Cancer (by no covariates) by Parish, Louisiana (1988-1993)

and was centered at three parishes (St. Mary, Iberia, and Assumption) of lower middle Louisiana. The first group in the secondary clusters was significant ($p = 0.001$), but the second group was insignificant ($p = 0.988$).

Considering the most likely cluster at both scales, the areas of spatial cluster by parish identified St. Bernard and Orleans parishes and those by census tract included 5 tracts in St. Bernard, 170 tracts in Orleans, and 74 tracts in Jefferson. The coincidence in clusters at the two levels could be easily expected because of the dominance of lung cancer deaths in parishes of southeastern Louisiana.

For the significant secondary cluster found at the parish level, there was an unexpectedly broad cluster in northern Louisiana. Its distribution was so general that most of them provided little additional information. But their existence indicated that while it is possible to represent the general location of a cluster, its exact boundaries remain uncertain.

The results of scan statistic by census tract (Figure 6.14) showed that there was a number of significant secondary clusters. To provide better understanding of the clusters of lung cancer (by no covariates), the crude death rates are shown in Figure 6.18. The geographic distribution of crude death rates (by only number of deaths and population) was relatively similar to that of clusters. In Figure 6.14, Clusters 2 was centered at 41 tracts, included 6 tracts of Bossier and 35 tracts of Caddo parishes. Clusters 3, 4, and 5 appeared at 12 tracts of Calcasieu, 39 tracts of East Baton Rouge, and 13 tracts of Ouachita, respectively. In the analysis and mapping of cancer data at the parish level, it was difficult to see important high death rates in some tracts. For

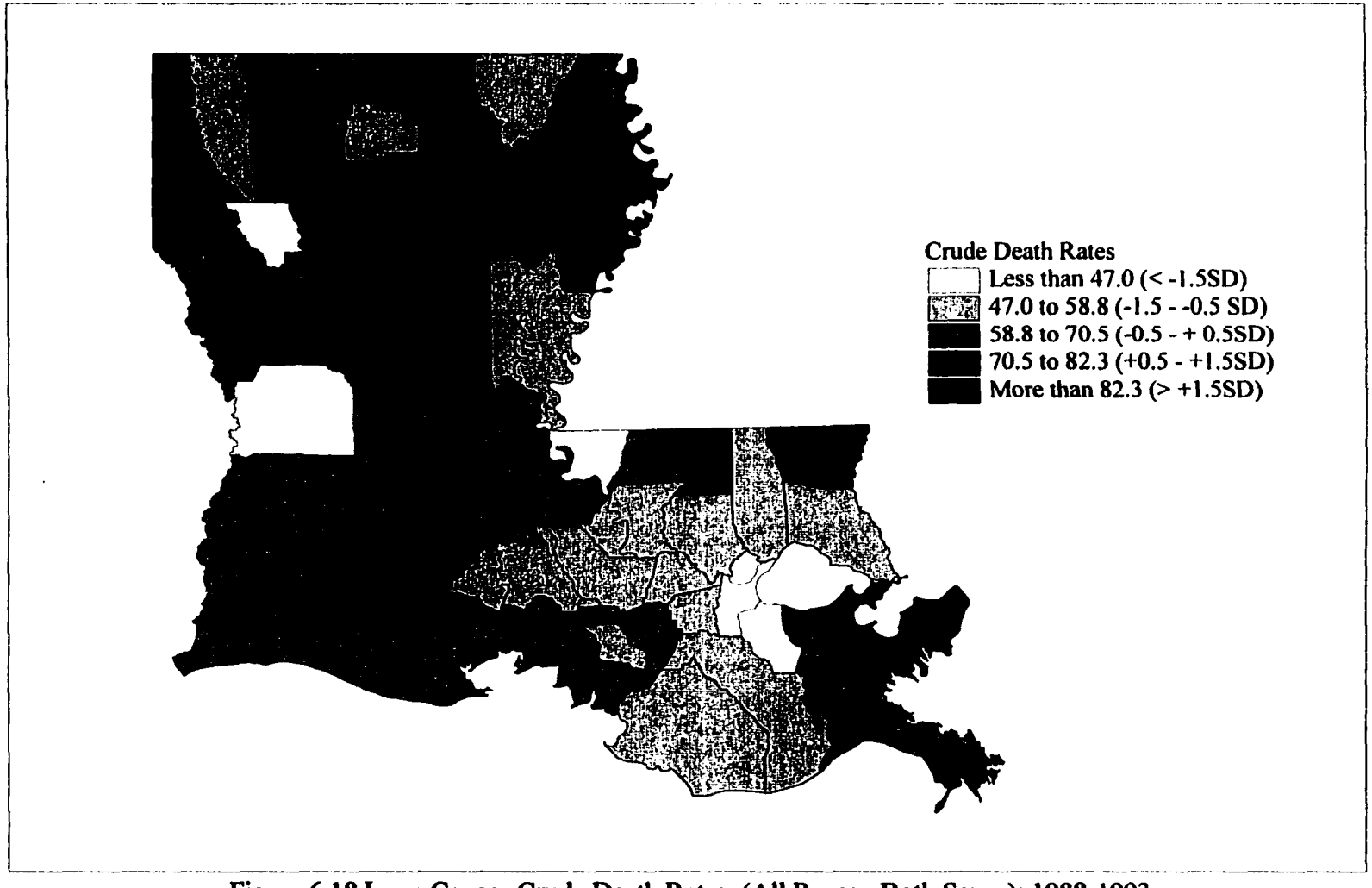

Figure 6.18 Lung Cancer Crude Death Rates (All Races, Both Sexes): 1988-1993
(Source: Centers for Disease Control and Prevention, 1999)

example, the cluster formed from tracts of Caddo and Bossier did not show high death rates when analyzed at the Bossier parish level.

Also, Clusters 6, 7, and, 8 appeared in Union (1 tract), Lafayette (7 tracts), and Evangeline (4 tracts). These clusters were significant at the level $p = 0.001$. Clusters 9, 10, and 11, which also were significant, included 2 tracts (Acadia), 6 tracts (Rapides), and 2 tracts (Vernon), with p-values of 0.004, 0.006, and 0.042, respectively. In case of cluster 6, it was difficult to consider it a cluster because this size was so small and this was single tract. Therefore, the spatial analysis at the small scale sometimes led to results that had not been revealed before. The findings here showed that scale and changes in scale affected the analysis and interpretation about geographic clusters of lung cancer. Changing the scale of data without first understanding their effects can result in the representation of processes or patterns that are different from those intended.

This study showed that to understand the cluster of lung cancer, empirical researchers are recommended to simultaneously analyze and represent data at multiple scales. To evaluate an area with high cancer mortality rates, and compare the results based on data at different scales, the spatial scan statistic was a useful statistical approach, with a significance test provided. However, more work is needed to identify and develop efficient spatial techniques for assessing and characterizing the scale effects.

6.2 Summary

Using the spatial scan statistic method, a large set of lung cancer deaths in Louisiana for the presence of geographical clusters at different scales (parishes and

census tracts) from 1988 to 1993 were tested. This analysis focused on geographic clusters with high rates, and the results of the analysis are as follows:

1. At the parish level, the results of spatial analysis for clusters with high lung cancer rates showed one significant most likely cluster and two insignificant secondary clusters. The most likely cluster could be characterized as "urbanization or industrialization" in southeastern Louisiana (St. Bernard, Plaquemines, and Jefferson parishes).

The clusters of lung cancer among white males, white females, and nonwhite females appeared in southeastern Louisiana whereas the cluster of nonwhite males appeared in southwestern Louisiana. It is unexpected that no significant clusters were detected among white males. In general, lung cancer among males and nonwhites had broader clusters with high rates than that among females and whites.

2. At the census tract level, the results of cluster detection of lung cancer deaths generated one significant most likely cluster and eleven significant secondary clusters. The most likely cluster appeared in 254 tracts along the Mississippi River in New Orleans region and the secondary clusters were widely distributed in Louisiana.

3. When the geographic clusters of lung cancer (by no covariates) were compared at the two different scales, the parish level generated one significant most likely cluster in St. Bernard and Orleans and the census tract level developed one significant most likely cluster along the Mississippi River in the New Orleans region. The most likely clusters at both scales were very similar, but the secondary clusters were dissimilar.

This analysis revealed that there was a statistically significant and geographically distinct cluster of lung cancer deaths in southeastern Louisiana from 1988 to 1993. It suggested that in order to investigate clusters of lung cancer, the spatial scan statistic is a useful statistical approach and the spatial analysis of cluster should be simultaneously performed at multiple scales, whenever possible.

CHAPTER 7

SUMMARY AND CONCLUSIONS

Spatial patterns of major cancer sites and their relationship to environmental factors in Louisiana were studied. The summary and conclusions are derived from the empirical results, focusing on the hypotheses presented in Chapter 1.

7.1 Hypothesis One

For hypothesis one (Cancer mortality rates such as breast, colon and rectum, lung, prostate, and stomach, are higher in South Louisiana than in the nation and the state), statistical mapping, significance test of rate differences, and factor analysis were used to analyze cancer mortality rates and trends in Louisiana compared with those in the nation from 1953 to 1987. The northeastern U.S. during 1953-1987 showed high cancer mortality rates whereas the mountain areas had relatively low rates. Of the 5 cancer sites, breast and colorectal cancer mortality rates developed similar distributions of high rates in the northeastern quadrant of the U.S. High lung cancer mortality rates prevailed in the southeastern counties, the West Coast, and the northeastern part of the country. High mortality rates for prostate cancer occurred in several counties of South Carolina, North Carolina, and Utah and those for stomach cancer occurred in the upper Midwest, northern New Mexico, and southeastern Louisiana. In particular, cancer mortality rates of most parishes in South Louisiana were significantly higher than those of the U.S. for cancers of all sites combined, lung and stomach of sex-race combined, lung among males, and stomach among nonwhite males.

Statistical maps of mortality rates were useful aids in describing and understanding spatial relationships. However, the classification methods used in

statistical mapping determine the appearance and message of a map and affect the visual interpretation. The standard deviation method used for this study has enhanced map presentation, comparison, and interpretation.

In factor analysis of cancer mortality rates in U.S. counties (Section 4.3) and Louisiana parishes (Section 5.2) from 1953 to 1987, major factors in the U.S. for 1953-1987 could be interpreted as "lung cancers for whites" (factor 1) and "cancers of breast and colorecutm for whites (factor 2) and nonwhites" (factor 3). However, within Louisiana, major cancer sites related to extracted factors were cancers of the digestive system (factor 1) and lung cancer for nonwhites (factor 2) during the two periods (1953-1977 and 1978-1987). The distributions of high scores for factor 1 between the two time periods have been extended from southeastern Louisiana to southern Louisiana, whereas those of factor 2 have been changed from southern Louisiana to southeastern Louisiana. In general, the geographical distributions of high scores of major factors were more prominent in southern Louisiana. The above findings indicated that South Louisiana had significantly higher cancer mortality rates for lung cancer among males and stomach cancer among nonwhite males, than the U.S. as the whole. Cancer of the digestive system and cancer among nonwhite in Louisiana were more serious problem than those in the U.S.

7.2 Hypothesis Two

Hypothesis two was to test if there were spatial clusterings of cancer deaths at Louisiana parish and census tract levels. The spatial autocorrelation and correlogram analyses were undertaken to examine the spatial-temporal patterns of cancer mortality rates in Louisiana for 1953-1987, 1953-1977, and 1978-1987. From 1953 to 1987,

cancers of all sites combined, lung, and stomach cancers exhibited high positive spatial autocorrelation, whereas cancers of breast, colorectum, and prostate exhibited low autocorrelation. Of all the correlograms, the shape of correlograms for lung cancer was the most distinctive and changed considerably through time.

A spatial scan statistic was used to test and compare the geographical clusters for lung cancer deaths at different levels from 1988 to 1993. At the parish level, the cluster detection results showed one significant most likely cluster in three parishes (St. Bernard, Plaquemines, and Jefferson parishes) at the tip of southeastern Louisiana. In the case of clusters of sex and race-specific lung cancer, one significant most likely cluster of lung cancer among white and nonwhite females centered in southeastern Louisiana, whereas that among nonwhite males appeared in southwestern Louisiana. However, among white males, no significant cluster was detected for lung cancer from 1988 to 1993.

At the census tract level, one significant most likely cluster of lung cancer appeared in tracts along the Mississippi River in the New Orleans region. Eleven significant secondary clusters were widely distributed in Louisiana. Comparing the geographic clusters of lung cancer deaths (by no covariates) between the two scales, the most likely clusters at both scales were very similar (St. Bernard and Orleans parishes at the parish level and a total of 254 census tracts in Jefferson, Orleans, and St. Bernard parishes at the census tract level), but the secondary clusters were dissimilar. In the case of the secondary cluster, there was one significant secondary cluster at the parish level and a number of secondary clusters at the census tract level. In sum, the cluster detection analysis showed that there was a statistically significant and geographically

distinct cluster of lung cancer deaths in southeastern Louisiana from 1988 to 1993. The most likely cluster of lung cancer could be characterized as "urbanization or industrialized areas in southeastern Louisiana."

A cluster is the occurrence of a greater than expected number of cases of a particular disease within a geographic area, a group of people, or a period of time (National Cancer Institute website at http://www.nci.nih.gov 1999). Some cancer clusters are not shown to be true clusters for the following reasons: many cancer clusters simply do not include enough cases to investigate and arrive at any conclusions; sometimes a cluster has enough cases, but a true statistical excess cannot be found; sometimes there is a true excess, but no explanation can be found. A random excess of cancer can occur but this does not necessarily mean that it can be linked to environmental or other factors. Also, because of the movement of population, it may be difficult to identify previous exposures. This study suggested that in order to investigate clusters of lung cancer, the spatial scan statistic is a useful statistical approach.

7.3 Hypothesis Three

For hypothesis three, which hypothesizes that cancer mortality patterns are associated with environmental variables, factor analysis and multiple regression were used. First, factor analysis was performed on 24 independent variables to determine the characteristics of environmental variables. The environmental variables were grouped into four factors: physical environment, chemical pollution, socio-economic status, and demographic resources.

The results of the stepwise regression analysis showed cancer mortality rates had a positive relationship with urban population and a negative relationship with persons employed in health and education services. Breast cancer mortality rates were found to be significantly positively related to urban population and education status. Colorectal cancer mortality rates were positively associated with drinking water (Mississippi River). In particular, lung cancer rates showed a positive relationship with persons employed in transportation and agricultural chemicals. In other words, lung cancer mortality rates were more closely related with occupational variables than any other cancer mortality rates and subject to a more potential health effect of exposure to toxic substances. Prostate and stomach cancer mortality rates were positively associated with nonwhite populations. Stomach cancer mortality rates showed a positive relationship with wetland areas. A better understanding of the characteristics of not only wetlands but also areas that include many wetlands is needed. This study provided additional information that was not visually evident from previous map and helped determine whether the geographic variation in cancer rates is related to urbanization, socioeconomic status, industrial exposures, or other indices.

Cancer is a group of diseases that occur when cells become abnormal and divide without control or order. This is usually not caused by a single factor, but is almost always caused by a combination of factors, including lifestyle, heredity, and environment, which interact in ways that are not yet fully understood. Today's cancer death rates reflect environmental conditions of the previous 10 to 30 years. This dissertation has contributed to our understanding of cancer and environment by spatial analysis identifying evidence of "cancer cluster." This study further asserts that

"environment" to cause cancer means the total (or accumulated) exposure of the individual to the internal and external world of physical, socio-economic, and mental factors, including genetic factors, in uncontrolled environment.

To minimize the risk of cancer in the environment, extensive research and cancer education are needed. Also, we need to recognize more about how politics (eg. ten year industrial property tax exemption for industrial support and environmental equity) and science (eg. occupied projects) shape the study of the causes of cancer and the policy to reduce such risk.

7.4 Future Research

First, this study used environmental data at the parish level because of limited data availability. Future research should be done with data at the finer level (tract or block) to better assess a causal relationship between cancer and environment. Moreover, the collected statistical data of cancer and environment were different from because they were different sources and by different statistic methods. Data quality and consistency is important issue in this type of study.

Second, when comparing the geographical cluster of lung cancer deaths at the parish and census tract levels, scale and changes in scale affect the analysis and interpretation about the geographic clusters of lung cancer. Changing the scale of data will result in the representation of processes or patterns that are different from those intended. Therefore, spatial analysis of cancer and environment should be done to study and compare the results at different spatial and temporal scales to reduce uncertainty.

Third, to find the causes of cancer attributed to environmental factors, correlation as well as case control studies will be needed. More disaggregate data on cigarette smoking, drinking, dietary factors, family history, and pollution over a long term are also needed.

Fourth, this study used GIS to store, manage, map, integrate, analyze, and visualize a large cancer and environment database. Most of the original data were not yet in directly usable form for GIS. A lot of time consuming database manipulation had to be made. For example, the geographic locations of some existing superfund sites were not clear. These sites were redigitized and the distance of 2 miles from the centroid of each superfund site was calculated by GIS software. Also, the area of wetlands of each parish were extracted and digitized from classified habitat image. Future study for environmental health should overcome current limitations of GIS data by standardizing data and making them more available so that they can be analyzed more quickly.

REFERENCES

Abler, R., Adams, J., and Gould, P. 1971. Spatial Organization: The Geographer's View of the World. Englewood Cliffs, New Jersey: Prentice Hall Inc.

Aldrich, T., Savitz, L., and Wittie, P. 1994. Evaluating Clusters of Adverse Health Outcomes. In Geographic Methods for Health Service Research, eds. T. Ricketts et al. :277-302. Maryland: University Press of America, Inc.

Alt, K.W. and Vach, W. 1991. The Reconstruction of Genetic Kinship in Prehistoric Burial Complexes-Problems and Statistics. In Classification, Data Analysis, and Knowledge Organization, eds. H.H. Back and P. Ihm. Berlin: Springer Verlag.

American Cancer Society. 1985, 1986, 1987, 1988, 1989, 1990, 1991, 1992, 1993, 1994, 1995, 1996, 1997, 1998, 1999. Cancer Facts and Figures. Atlanta: American Cancer Society, Inc.

American Cancer Society. 1990a. The Louisiana Cancer Problem: Questions and Answers. Louisiana: American Cancer Society, Inc.

American Cancer Society. 1992a. Prevention and Control. Cancer Risk Report. Atlanta: American Cancer Society, Inc.

Anselin, L. and Getis, A. 1992. Spatial Statistical Analysis and Geographic Information Systems. Annals of Regional Science 26:19-33.

Anselin, L., Dodson, R.F., Hudak, S. 1993. Linking GIS and Spatial Data Analysis in Practice. Geographical Systems 1:3-23.

Armstrong, R.W. 1969. Standardized Class Intervals and Rate Computation in Statistical Maps of Mortality. Annals of the Association of American Geographers 59:382-390.

Armstrong, R.W. 1972. Computers and Mapping in Medical Geography. In Medical Geography: Techniques and Field Studies, ed. N.D. MaGlashan, :69-88. London: Methuen.

Armstong, R.W. 1976. The Geography of Specific Environments of Patients and Non-Patients in Cancer Studies, with a Malaysian Example. Economic Geography 52:161-170.

Armstrong, R.W. 1979. Self-Specific Environments Associated with Nasopharyngeal Carcinoma in Selangor, Malaysia. Social Science and Medicine 120D:149-156.

Aronchick, Judith M. 1990. Lung Cancer: Epidemiology and Risk Factors. Seminars in Roentgenology 25(1): 5-11.

Babin, E. 1979. United States Cancer Mortality Regions: 1959-1969. Social Science and Medicine 13D:39-43.

Babin, E. 1986. United States Chapter 3. In Global Geocancerology, ed. G.M. Howe, pp. 61-89. New York: Churchill Livingstone.

Baker, R.D. 1996. Testing for Space-Time Clusters of Unknown Size. Journal of Applied Statistics 23:543-554.

Bennett, D. 1991. Expiation in Medical Geography: Evidence and Epistemology. Social Science and Medicine 33(4):339-346.

Besag, J. and Newell, J. 1991. The Detection of Clusters in Rare Diseases. Journal of the Royal Statistical Society, Series A 154:143-155.

Bithell, J. F. 1995. The Choice of Test for Detecting Raised Disease Risk Near a Point Source. Statistics in Medicine 14:2309-2322.

Blot, W.J. and Fraumeni, J.F(Jr). 1976. Geographic Patterns of Lung Cancer: Industrial Correlations. American Journal of Epidemiology 103:539-550.

Blot, W.J., Fraumeni, J.F(Jr)., and Stone, B.J. 1977. Geographic Patterns of Breast Cancer in the United States. Journal of National Cancer Institute 59:1407-1411.

Blot, W.J. and Fraumeni, J.F(Jr). 1978. Geographic Patterns of Bladder Cancer in the United States. Journal of National Cancer Institute 61:1017-1023.

Blot, W.J., Fraumeni, J.F(Jr), and Stone, B.J. 1978. Geographic Correlates of Pancreas Cancer in the United States. Cancer 42:373-380.

Blot, W.J., Fraumeni, J.F., Mason, T.J., and Hoover, R.N. 1979. Developing Clues to Environmental Cancer: A Stepwise Approach with the Use of Cancer Mortality Data. Environmental Health Perspectives 32:53-58.

Blot, W.J., Fraumeni, J.F(Jr)., and Mason, T.J. 1982. Changing Patterns of Lung Cancer in the United States. American Journal of Epidemiology 115:664-670.

Boring, C.C., Squires, T.S., and Tong, T. 1993. Cancer Statistics. Atlanta: American Cancer Society Inc.

Briggs, R. and Leonard IV, W.A. 1977. Mortality and Ecological Structure: A Canonical Approach. Social Science and Medicine 11:757-762.

Brown and Donnelly. 1988. An Estimation of the Risk Associated with the Organic Constituents of Hazardous and Municipal Waste Landfill Leachates. Hazardous Waste and Hazardous Materials 5(1).

Bruncher, C.R., Kuzma, R.J., and Forcade, C.M. 1977. Drinking Water as an Epidemiologic Risk Factor for Cancer. In Origins of Human Cancer: Book A. Incidence of Cancer in humans, eds. H.H. Hiatt et al., PP.347-366. Cold Spring Harbor, Cold Spring Harbor Laboratory.

Burbank, F. 1971. Patterns in Cancer Mortality in the United States: 1950-1967, NCI Monograph 33. Washington, D.C.: U.S. Government Printing Office.

Cantor K.P., Hoover, R., Hartge, P., et al. 1987. Bladder Cancer Drinking Water Source, and Tap Water Consumption: A Case-Control Study. Journal of National Cancer Institute 79:1269-1279.

Centers for Disease Control, Nutrition program. 1972. Ten-state Nutritive Survey. 1968-1970. U.S. Department of Health, Education and Welfare, Publications (HSM) 72-8129, 72-8130, 72-8131, 72-8132, 72-8133.

Chen, V.W., Correa, P., and Diego, Z.E. 1984. Lung Cancer: Leading Cause of Cancer Deaths in Louisiana White Females. Journal of National Cancer Institute 72:1-2.

Chen, V.W., Correa, P., Craig, J.F., and Fontham, E.T.H. 1988. Cancer Survival Poorer in Louisiana? Journal of the Louisiana state Medical Society 140(4):20-28.

Chen, V.W., Craig, J.F., Fontham, E., Lin, Y. and Correa, P. 1990. Cancer in Louisiana, vol VI: Cancer Incidence in South Louisiana 1983-1986. New Orleans: Louisiana Tumor Registry Office of Public Health.

Chen, V.W., Fontham, E., Groves, F.D., Craig, J.F., and Correa, P. 1991. Cancer in Louisiana, Vol VII: Cancer Incidence in South Louisiana 1983-1986. New Orleans: Louisiana Tumor Registry Office of Public Health.

Chen, V.W., Fontham, E.T.H., Graig, J.F., et al. 1992a. Cancer in South Louisiana, Part 1: Tobacco-related Cancers. Journal of the Louisiana State Medical Society - Special Issue 144(4):149-155.

Chen, V.W., Groves, F.D. Fontham., E.T.H., Craig, J.F., et al. 1992b. Cancer in South Louisiana, Part III: Cancer of the Breast and the Reproductive System. Journal of the Louisiana State Medical Society 144:171-177.

Chen, V.W., Wu, X., Andrews, P.A., and Fontham, E.T.H. 1994. Advanced Stage at Diagnosis: An Explanation for Higher Than Expected Cancer Death Rates in Louisiana? Journal of the Louisiana State Medical Society 146:137-145.

Chen, V.W., Correa, C.N., Wu, X.C., Andrews, P.A., et al. 1996. Cancer Incidence in Louisiana 1988-1992 VIII(1-9). New Orleans: Epidemiology section, Louisiana state University Medical center.

Chen, V.W., Wu, X., Andrews, P.A., and Lucas., H.F. 1997. Highlights of Cancer Incidence in Louisiana, 1988-1992. Journal of the Louisiana State Medical Society 149:199-124.

Chen, V.W., Wu, X., Andrews, P.A., et al. 1998. Cancer Incidence in the Industrial Corridor: An Update. Journal of the Louisiana State Medical Society 150:158-167.

Chiang, C.L. 1961. Standard Error of the Age-adjusted Death Rate. Vital Statistics Selected Reports 47(9). Washington.

Chou, Y.H. 1991. Map Resolution and Spatial Autocorrelation. Geographic Analysis 23(3):228-246.

Chou, Y.H. 1993. Critical Issues in the Evaluation of Spatial Autocorrelation. In Lecture Notes in Computer Science, Spatial Information Theory, eds. A.U. Frank and I. Campari, pp.716:421-433. New York, Berlin: Springer-verlag.

Choynowski, M. 1959. Maps Based on Probabilities. Journal of the American Statistical Association 54:385-388.

Clayton, D.G., Bernardinelli, L., and Montomoli, C. 1993. Spatial Correlation in Ecological Analysis. International Journal of Epidemiology 22(6):1193-1202.

Cleek, R.K. 1979. Cancers and the Environment: The Effect of Scale. Social Science Medicine. 13D:241-247.

Cleverley, W.O.and Nutt , P.C. 1984. The Decision Process Used for Hospital Bond Rating-and Its Implication. Health Services Research 19(5):615-637.

Cliff, A.W. and Haggett, P. 1988. Atlas of the Distribution of Diseases, Analytical Approaches to Epidemiological Data. Oxford: Basil Blackwell Ltd.

Cliff, A.D., Haggett, P, and Ord, J.K. 1975. Elements of Spatial Structure: A Quantitative Approach. Cambridge: Cambridge University Press.

Cliff, A.D. and Ord, J.K. 1973. Spatial Autocorrelaation: New York: Methuen, Inc

Cliff, A.D. and Ord, J.K. 1981. Spatial Processes: Models and Applications. London: Pion.

Clipp, A (ed.). 1994. The Louisiana Legislative Briefing Book:1994 / The Environment. The Louisiana Environmental Action Network.

Coffer, J.V. and Patrick, L.L. 1981. Disease and Ethnicity in an Urban Environment. Annals of Association of American Geographers 71(1):40-82.

Correa, P., and Johnson, W.D. 1983. Cancer and Lifestyle in Louisiana. Journal of Louisiana State Medical Society 135(3):4-6.

Correa, P., Chen, V.W., Craig, J., et al. 1983a. Cancer in Louisiana Vol I. Baton Rouge: Louisiana Division of Administration, Administrative Services :1-248.

Correa, P., Pickle, L.W., Fontham, E., Lin, Y.P., and Haenszel, W. 1983b. Passive Smoking and Lung Cancer. Lancet 8350:595-597.

Correa, P., Pickle, L.W., and Fontham, E.T.H. 1984a. Cancer in Louisiana: Case-control Studies of Environmental Determinants in High Risk Parishes: Final Report. Proceedings of the Third NCI/EPA/NIOSH Collaborative Workshop: Progress on Joint Environmental and Occupational Cancer Studies:727-751.

Correa, P., Pickle, L.W., Fontham, E., Dalager, N., and Lin, Y.P. 1984b. The Causes of Lung Cancer in Louisiana. In Lung Cancer Causes and Prevention, eds. M. Mizelli and P. Correa, pp. 73-82. Fl: Verlag Chemie International, Inc.

Correa, P., Chen, V.W., Craig, J., et al. 1985a. Cancer in Louisiana Vol II. Baton Rouge: Louisiana Division of Administration, Administrative Services: 1-47.

Correa, P., Fontham, E.T.H., Pickle, L.W., Chen, V.W., Lin, Y., and Haenszel, W. 1985b. Dietary Determinants of Gastric Cancer in South Louisiana Inhabitants. Journal of National Cancer Institute 75:645-654.

Correa, P., Chen, V.W., Craig, J., et al. 1986. Cancer in Louisiana Vol III. Baton Rouge: Louisiana Division of Administration, administrative Services:1-122.

Correa, P., Chen, V.W., Craig, J., et al. 1987. Cancer in Louisiana Vol IV. Baton Rouge: Louisiana Division of Administration, Administrative Services:1-150.

Correa, P., Fontham, E.T.H., Pickle, L.W., and Chen, V.W. 1988. Diet, Nutrition, and Cancer. Journal of the Louisiana State Medical Society 140:43-50.

Correa, P., Chen, V.W., Craig, J., et al. 1989. Cancer in Louisiana Vol V. Baton Rouge: Louisiana Division of Administration, Administrative Services:1-93.

Costner P. and Thornton J. 1989. We All Live Downstream: The Mississippi River and the National Toxics Crisis. Greenpeace USA.

Cowardin, L.M. et al. 1992. Classification of Wetlands and Deepwater Habitats of the United States. U.S. Department of the Interior: Fish and Wildlife Service.

Craig, J.F., Chen, V.W., et al. 1990. Cancer Mortality in the 30s Foretells Cancer of the 80s in Louisiana. Journal of the Louisiana state Medical Society 142:43-48.

Craig, J.F., Chen, V.W., Fontham, E.T.H., et al. 1991. Cancer Control in Louisiana: Assessment of Needs and Development of a Plan. Journal of Louisiana State Medical Society 143:41-46.

Cuzick, J. and Edwards, R. 1990. Spatial Clustering for Inhomogeneous Population. Journal of the Royal Society, Series B 52:73-104.

Davis, D.L. and Hoel, D. (eds.). 1990. Trends in Cancer Mortality in Industrial Countries (Annals of the New York Academy of Sciences 609). New York: New York Academy of Sciences.

de Lepper, Marion J.C. and Scholten, Henk J., and Stern, Richard M.(eds.). 1995. The Added Value of Geographical Information Systems in Public and Environmental Health. Boston: Academic Publishers.

De Savigny, D. and Wijeyaratne, P. (eds.). 1994. GIS for Health and the Environment. Proceedings of an International Workshop, Colombo, Sri Lanka. Ottawa: International development Research Center.

Dent, O.F. and Goulston, K.J. 1980. Public Awareness of Cancer: a Canberra Survey. Cancer Forum 19:35-37.

Dent, O.F. and Goulston, K.J. 1982. A Short Scale of Cancer Knowledge and Some Socio-demographic Correlates. Social Science and Medicine 16:235-245.

DeRouen, T.A. and Diem, J.E. 1975. The New Orleans Drinking Water Controversy: A Statistical Perspective. American Journal of Public health 65:1060-1062.

DeRouen, T.A. and Diem, J.E. 1977. Relationships between Cancer Mortality in Louisiana Drinking-water Source and Other Possible Causative Agents. In Origins of Human Cancer: Book A. Incidence of Cancer in Humans, eds. H.H. Hiatt et al., pp. 331-345. Cold Spring Harbor: Cold Spring Harbor Laboratory.

Devesa, S.S. and Silverman, D.T. 1978. Cancer Incidence and Mortality Trends in the United States: 1935-1974. Journal of National Cancer Institute 60(3):545-571.

Diggle, P.J. and Chetwynd, A.D. 1991. Second-Order Analysis of Spatial Clustering for Inhomogeneous Population. Biometrics 47:1155-1163.

Diggle, P.J., Chetwynd, A.G., Morris, S.E., et al. 1995. Second-Order Analysis of Space-Time Clustering. Statistical Methods in Medical Research 4:124-136.

Djoundourian, Salpie Sarkis. 1993. The Distribution of Toxic Sources in Louisiana: An Analysis of Community Socioeconomic Characteristics and Industry Behavior. Louisiana State University: Unpublished Doctoral Dissertation.

Doll, R. 1969. The Geographical Distribution of Cancer. British Journal of Cancer 23:1-18

Doll, R. 1990. Health and the Environment in the 1990s. American Journal of Public Health 82:933-941.

Doll, R. and Peto, R. 1981. The Causes of Cancer: Quantitative Estimates of Avoidable Risks of Cancer in the United States Today. New York: Oxford University Press.

Douven, W. and Scholten, H.J. 1995. Spatial Analysis in Health Research. In The Added Value of Geographical Information Systems in Public and Environmental Health, eds. M.J.C. de Lepper, H.J. Scholten, and R.M. Stern, pp.117-134. Boston: Kluwer Academic Publishers.

Dunham, L.J. and Bailar, J.C. 1968. World Maps of Cancer Mortality Rates and Frequency Ratio. Journal of the National Cancer Institute 41:155-203.

Dunham, L.J and Chiazze, L. Jr., Stewart, H., and Young J. 1968. Rates, Interview, and Pathology Study of Cancer of the Urinary Bladder in New Orleans, Louisiana. Journal of National Cancer Institute 41:683-709.

Earle, C. 1979. Environment, Disease, and Mortality in Early Virginia. Journal of Historical Geography 5:369-390.

Elaimy, W.M. 1990. Future Directions in Cancer Control for the Louisiana Cancer Consortium. Journal of the Louisiana State Medical Society 142:33-34.

Elliott, P, Cuzick, D.E, and Stern, R. 1997. Geographical & Environmental Epidemiology: Methods for Small-Area Studies. New York: Oxford Medical Publications.

Environmental Protection Agency. 1984. The Magnitude and Nature of the Air Toxic Problem in the United States. Washington, D.C.: US Government Printing Office.

Enstrom, J.E. 1978. Cancer and Total Mortality among Active Mormons. Cancer 42: 1943-1971.

Environmental Systems Research Institute, Inc. 1996. ArcView GIS: Using ArcView GIS.

Esteve, J., Benhamou, E., and Raymond, L. 1994. Statistical Methods in Cancer Research: Descriptive Epidemiology. Lyon: International Agency for Research on Cancer

Eyles, J. 1983. Man, Disease, and Environmental Associations: from Medical of Medicine and Health. New York: St. Martin's Press.

Falk, R.T., Pickle, L.W., Fontham, E.T., et al. 1988. Life-Style Risk Factors for Pancreatic Cancer in Louisiana: A Case- Control Study. American Journal of Epidemiology 128(2):324-336.

Fan, M., Lam, N.S.-N., and Liu, K.B. 1993. Using GIS to Model the Spatial Spread of the AIDS Epidemic in the United States. Technical Papers ACSM/ASPRS Annual Convention:186-195.

Fandrich, J. and Zwick, P. 1993. Using a GIS to Evaluate Radon Potential and its Effect on Housing. Abstracts 89th Annual Meeting of the Association of American Geographers 68. Washington, D.C.:Association of American Geographers.

Fausett, L. 1994. Fundamentals of Neural Networks: Architectures, Algorithm, and Application. Prentice-Hall, Inc.

Feng, W. 1998. Projecting the AIDS Epidemic in the United States Using an Artificial Neural Network Model. An Unpublished Thesis: Department of Geography and Anthropology of Louisiana State University, Baton Rouge.

Ference, L., Chiazze, L.Jr., and Wolf, P. 1987. A Cohort Mortality Study of Four Plants of the Allied Corporation. Journal of Occupational Medicine 29:839-841.

Fischer, M. and Nijkamp, P. 1992. Geographic Information Systems and Spatial Analysis. Annals of Regional Science 26:3-17.

Fischer, M. and Nijkamp, P. 1993. Geographic Information Systems, Spatial Modelling, and Policy Evaluation. New York: Springer-Verlag.

Fitzpatrick-Lins, K. et al. 1990. Radon Potential Defined by Exploratory Data Analysis and Geographic Information Systems. Selected Papers in the Applied Computer Sciences 1990. ed. D.A. Wiltshire, U.S. Geological Survey Bulletin 1908: E1-E-10.

Fontham, E.T.H., Correa P., Chen, V.W., Craig J.F., Pickel L.W., et al. 1988a. Tobacco and Cancer. Journal of the Louisiana State Medical Society 140:29-40.

Fontham, E.T.H., Pickle, L., et al. 1988b. Dietary Vitamins A and C and Lung Cancer Risk in Louisiana. Cancer 62(10):2267-2273.

Fontham, E.T.H., Chen, V.W., Craig, J.F., et al. 1992. Cancer in South Louisiana. Part II: Gastrointestinal Tract Cancers. Journal of Louisiana State Medical Society - Special Issue 144(4):163-166.

Fontham, E.T.H., Correa, P., Chen, V.W., et al. 1993. Passive Smoking and Lung Cancer. Journal of Louisiana State Medical Society 145:132-136.

Forster, F. 1966. Use of a Demographic Base Map in the Presentation of Areal Data in Epidemiology. British Journal of Preventative and Social Medicine 20:165-171.

Garfinkel, L. 1990. Guest Editorial - The Environment and Cancer: Putting the Risks into Perspective. Ca-A Cancer Journal for Clinicians 40(5):261-263.

Gatrell, A.C. and Dunn, C.E. 1995. Geographical Information Systems and Spatial Epidemiology: Modelling the Possible Association Between Cancer of the Larynx and Incineration in North-West England. In The Added Value of Geographical Information Systems in Public and Environmental Health, eds. M.J.C. de Lepper, H.J. Scholten, and R.M. Stern: 215-235. Boston: Kluwer Academic Publishers.

Gesler, W. 1986. The Use of Spatial Analysis in Medical Geography: A Review. Social Science and Medicine 23(10):963-973.

Glick, B.J. 1977. Environment and Disease: A Spatial Analytic Model and Application to Cancer in Pennsylvania. New York: Cornell University Press.

Glick, B.J. 1979a. Distance Relationships in Theoretical Models of Carcinogenesis. Social Science and Medicine 13D:253-256.

Glick, B.J. 1979b. The Spatial Autocorrelation of Cancer Mortality. Social Science and Medicine 13D:123-130.

Glick, B.J. 1982. The Spatial Organization of Cancer Mortality. Annals of the Association of American Geographers 72:471.

Goddard, J. and Kirby, A. 1976. An Introduction to Factor Analysis. Catmog 7.

Goins, C.R. and Caldwell, J.M. 1995. Historical Atlas of Louisiana. Norman: University of Oklahoma.

Goldman, B.A. 1991. The Truth About Where You Live: An Atlas for Action on Toxins and Mortality. New York: Random House, Inc. (Times Books).

Goldstein, B.D. 1992. Cancer and the Environment. Science 255:904.

Goodall, B. 1987. The Facts on File Dictionary of Human Geography. New York: Facts on Files, Inc.

Goodchild, M.F. 1986. Spatial Autocorrelation. CATMOG (Concepts and Techniques in Modern Geography) 47. Norwich, England: Geo Books.

Goodchild, M.F. 1987. A Spatial Analytical Perspective on Geographical Information Systems. International Journal of Geographical Information Systems 1(4):327-334.

Goodchild, M.F. 1992. Integrating GIS and Spatial Data Analysis: Problems and Possibilities. International Journal of Geographical Information Systems 6(5):407-423.

Goodchild, M.F. and Brusegard, D. 1989. Spatial Analysis Using GIS: Seminal Workbook. National Center for Geographic Information Environmental Health Planning.

Gottlieb, M.S. 1980. Lung Cancer and the Petroleum Industry in Louisiana. Journal of Occupational Medicine. 22:384-388.

Gottlieb, M.S. and Stedman, R. 1979. Lung Cancer in Shipbuilding and Related Industries in Louisiana. Southern Medical Journal 72:1099-1101.

Gottlieb, M.S., Pickle, L.W., Blot, W.J., Fraumeni, J.F. 1979. Lung Cancer in Louisiana: Death Certificate Analysis. Journal National Cancer Institute 63:1144-1137.

Gottlieb, M.S. and Carr, J.K. 1981. Mortality Studies on Lung, Pancreas, Esophageal, and Other Cancers in Louisiana. In Banbury Report 9: Quantification of Occupational Cancer, eds. R.Peto, et al.:195-204. Cold Spring Harbor: Cold spring Harbor Laboratory.

Gottlieb, M.S. and Pickle, L.W. 1981. Bladder Cancer Mortality in Louisiana. The Journal of the Louisiana State Medical Society 133:6-9.

Gottlieb, M.S., Carr, J.K., and Morris, D.T. 1981. Cancer and Drinking Water in Louisiana: Colon and Rectum. International Journal of Epidemiology 10(2):117-125.

Gottlieb, M.S. and Carr, J.K. 1982a. Case-control Cancer Mortality Study and Chlorination of Drinking Water in Louisiana. Environmental Health Perspectives 46:169-177.

Gottlieb, M.S. and Carr, J.K. 1982b. Drinking Water and Cancer in Louisiana: A Retrospective Mortality Study. American Journal of Epidemiology 116:169-177.

Gottlieb, M.S., Shear, C., and Seale, D. 1982c. Lung Cancer Mortality and Residential Proximity to Industry. Environmental Health Perspectives 45:157-164.

Gover, M. 1940. Cancer Mortality in the United States, part III. Geographic Variation in Recorded Cancer Mortality for Detailed Sites for an Average of the Years 1930-1932. Public Health Bulletin No. 757. Us Public Health service.

Green, N.P.A. 1987. Teach Yourself Geographical Information Systems: The Design, Creation and Use of Demonstrators and Tutors. International Journal of Geographic Information Systems 1(3):279-290.

Greenberg, M.R., Preuss, P.W., and Anderson, R. 1980. Clues for Case Control Studies of Cancer in the Northeast Urban Corridor. Social and Medicine 14(D):37-41.

Greenberg, M.R. 1983. Urbanization and Cancer Mortality: The United States Experience, 1950-1975. New York: Oxford University Press.

Greenberg, M. R. 1984. Sunbelt, Frostbelt and Public Health. Society 21(July-August): 68-75.

Greenwald, P. 1982. Prostate. In Cancer Epidemiology and Prevention, eds. D. Schottenfeld and J.F. Jr. Fraumeni:938-946. Philadelphia: W.B. Saunders.

Griffiths, M. 1971. A Geographical Study of Mortality in an Urban Area. Urban Studies 8:111-120.

Griffiths, William E., R. Carter Hill, and George G. Judge. 1993. Learning and Practicing Econometrics. New York: John Wiley & Sons

Grimson, R.C. 1991. A Versatile Test for Clustering and a Proximity Analysis of Neurons. Methods of Information in Medicine 30:299-303.

Groves, F.D., Griener, E.D., and Fontham, E.T.H. 1992. Lung Cancer Rates and Risk Factors in St. Bernard Parish, Louisiana: An Ecologic Study. Journal of Louisiana State Medical Society 144:271-277.

Groves, F.D., Andrews, P.A., Chen, V.W., Fontham, E.T.H., and Correa, P. 1996. Is There a 'Cancer Corridor' in Louisiana. Journal of Louisiana State Medical Society 148(4):155-165.

Haber, M.H. and Lipkovic, P. 1970. Thyroid Cancer in Hawaii. Cancer 25:1224-1227.

Haenszel, W. 1960. Variations in Skin Cancer Incidence within the United states. In The First International Conference on the Biology of Cutaneous Cancer, ed. F. Urbach. Bethesda: National Cancer

Haggett, P.A., Cliff, D., and Frey, A. 1977. Locational Analysis in Human Geography. New York: John Wiely & Sons.

Haining, R. 1990. Spatial Data Analysis in the Social and Environmental Sciences. New York: Cambridge University Press.

Hanis, N., Holmes, T., Shallenberger, L., and Jones, K. 1982. Epidemiologic Study of Refinery and Chemical Plant Workers. Journal of Occupational Medicine 24:203-212.

Harris, R.H. 1974. The Implication of Cancer-Causing Substances in Mississippi River Water. Washington, D.C.: Environmental Defense Fund.

Herman, D. 1981. Cohort Mortality Study of the Scotts Bluff/Baton Rouge Uniroyal Plant. Chapel Hill, North Carolina: University of North Carolina.

Herron, J. 1969. The Geographical Distribution of Malignant Melanoma in Queensland. Medicine Journal of Australia 2:892-904.

Heywood, I. 1990. Geographic Information Systems in the Social Sciences. Environment and Planning A 22:849-854.

Hiatt, H.H., Watson, J.D., and Winsten, J.A. 1977. Origins of Human Cancer: Incidence of Cancer in Humans. Cold Spring Harbor: Cold Spring Harbor Laboratory.

Higginson, J. 1988. Changing Concepts in Cancer Prevention: Limitations and Implications for Future Research in Environmental Carcinogenesis. Cancer Research 48:1381-1389.

Higginson, J. 1993. Environmental Carcinogenesis. Cancer Supplement August 1, 72(3):971-976.

Higginson, J., Muir, C.S., and Munoz, N. 1992. Human Cancer: Epidemiology and Environmental Causes. Cambridge: Cambridge University Press.

Hjalmars, U., Kulldorff, M., Gustafsson, G., and Nagarwalla, N. 1996. Childhood leukemia in Sweden: Using GIS and a spatial scan statistic for cluster detection. Statistics in Medicine 15:707-715.

Holmes, T. and Buffler, P., Holguin, A., et al. 1986. A Mortality Study of Employees at a Synthetic Rubber Manufacturing Plant. American Journal of Industrial Medicine 9:355-362.

Hoover, R. and Fraumeni, Jr.F. 1975. Cancer Mortality in U.S. Counties with Chemical Industries. Environmental Research 9:196-207.

Hopps, H.C. 1969. Computer-produced Distribution Maps of Disease. Annals of the N.Y. Academy of Science Vol. CLXI.

Howe, G.H. 1960. The Geographical Distribution of Cancer Mortality in Wales 1947-53. Transactions of the Institute of British Geographers 28:199-214.

Howe, G.H. 1972. Man, Environment, and Disease in Britain. New York. Medicine 12d:7-15.

Howe, G.H. 1981. Mortality from Selected Malignant Neoplasms in the British Isles: The Spatial Perspective. Social Science and Medicine 15D:199-211.

Howe, G.H. 1986a. Does It Matter Where I Live?. Transactions of the Institute of British Geographers, New Series 6:387-414.

Howe, G.H. 1986b. Global Geocancerology. New York: Churchill Livingstone

Howe, H.L. and Lehnherr, M. (eds.). 1996. Cancer Incidence in North America, 1988-1992. California: North American Association of Central cancer Registries.

Hughes, J., Weill, H., and Hammad, Y. 1987. Mortality of Workers Employed in Two Asbestos Cement Manufacturing Plants. British Journal of Industrial Medicine 44:161-174.

Hunter, D.J. and Willett, W.C. 1996. Nutrition and Breast Cancer. Cancer Causes Control 7:56-58.

Inaba, Y., Yanai, H. and Takagi, H., et al. 1981. A Study of the Geographical Patterns of Cancer Mortality for Selected Sites by Means of Factor Analysis. Social Science and Medicine 15D:233-244.

Institute for Environmental Issues and Policy Assessment. 1994. RiverSentinel 1994: Environmental, Health and Social Trends in the Lower Mississippi River Corridor. Baton Rouge: Southern University.

Institute for Environmental Issues and Policy Assessment. 1995. RiverSentinel 1995: People Who Can Be Harmed By Toxic Releases. Baton Rouge: Southern University.

Jacquez, G.M. 1996. A *K* Nearest Neighbor Test for Space-Time Interaction. Statistics in Medicine 15:1935-1949.

Jakab, S. et al. 1971. Possible Relationships Between Geographical Environmental Factors and the Frequency of Gastric Cancer in East Transylvania. Geographic Medica Vol. 2.

Jones, K., Paul, K., et al. 1987. Health, Disease, and Society: An Introduction to Medical Geography. London and New York: Roultedge & Kegan Paul.

Joseph, A.E. and Phillips, D.R. 1984. Accessibility and Utilization: Geographical Perspectives on Health Care Delivery. New York: Harper & Row.

Kehris, E. 1990. Spatial Autocorrelation Statistics in ARC/INFO. North West Research. Laboratory Research Report No.16. Lancaster University.

Kennaway E.L. 1944. Cancer of the Liver in the Negro in Africa and America. Cancer Research 4:571.

Kennedy, U. 1988. A Geographic Regression Model for Medical Statistics. Social Science and Medicine 23(1):119-129.

Kim, S.Y. 1984. Past and Future of Medical Geography. Sangmong Geography 1:51-56.

Knox, G. 1964. The Detection of Space-Time Interactions. Applied Statistics 13:25-29.

Knox, E.G. and Gilman, E. 1992. Leukaemia Clusters in Great Britain. Journal of Epidemiology and Community Health 46:556-576.

Knutson, Ronald D., J. B. Penn, and Barry L. Flinchbaugh (4^{th} eds.). 1998. Agricultural and Food Policy. New Jersey: Princeton Hall.

Kolonel, L.N., Hankin, J.H., Lee, J. et al. 1981. Nutrient Intakes in Relation to Cancer Incidence in Hawaii. British Journal of Cancer 44:332-339.

Kolonel, L.N. 1996. Nutrition and Prostate Cancer. Cancer Causes Control 7:83-94.

Kono, S. and Hirohata, T. 1996. Nutrition and Stomach Cancer. Cancer Causes Control 7:41-55.

Kraybill, H.F. 1977. Global Distribution of Carcinogenic Pollutants in Water. Annals of New York Academy of Science 298:80-89.

Kuehn, R.R.(ed.) 1991. Citizen's Guide to Environmental Activism in Louisiana. Tulane Environmental La Clinic.

Kulldorff, M. 1997. A Spatial Scan Statistic. Communications in Statistics: Theory and Methods 26:1481-1496.

Kulldorff, M. and Hjalmars, U. 1999. The Knox Method and Other Tests for Space Time Interaction. Biometrics. In Press.

Kulldorff, M. and Nagarwalla, N. 1995. Spatial Disease Clusters: Detection and Inference. Statistics in Medicine 14:799-810.

Kulldorff, M., Feuer, E.J., Miller, B.A., and Freedman, L.S. 1997. Breast Cancer in Northeastern United States: a Geographical Analysis. American Journal of Epidemiology 146:161-170.

Kuzma, R.J., Kuzma, C.M., and Buncher, C.R. 1979. Ohio Drinking Water Source and Cancer Rates. American Journal of Public Health 67:725-729.

Lam, N.S.-N 1986. Geographical Patterns of Cancer Mortality in China. Social Science and Medicine 23(3):241-247.

Lam, N.S.-N. and De Cola, L. (eds.). 1993. Chapter 1: Introduction to Fractals in Geography. In Fractals in Geography, eds. N.S.- N. Lam and L. De. New Jersey: PTR Prentice-Hall.

Lam, N.S.-N., Qiu, H.-L., Zhao, R, and Jiang, N. 1993. Chapter 12: A Fractal Analysis of Cancer Mortality Patterns in China. In Fractals in Geography, eds. N.S.-N. Lam and L. De Cola, New Jersey: PTR Prentice-Hall.

Lam, N.S.-N, and Liu, K.-B. 1994. Spread of AIDS in Rural America, 1982-1990. Journal of Acquired Immune Deficiency Syndromes 7:485-490.

Lam, N.S.-N, Fan, M., and Liu, K.-B. 1996. Spatial-Temporal Spread of the AIDS Epidemic, 1982-1990: A Correlogram Analysis of Four Regions of the United States. Geographical Analysis 28(2): 93-107.

Lawson, A.B. 1993. On the Analysis of Mortality Events Associated with a Prespecified Fixed Point. Journal of the Royal Statistical Society Series A, 156:363-377.

Learmonth, A.T.A. 1972. Atlases in Medical Geography 1950-1970: A Review. In Medical Geography: Techniques and Field Studies, ed. N.D. McGlashan:133-152. London: Methuen & Co Ltd.

Lenihan, J. and Fletcher, W.W. 1976. Health and the Environment. New York: Academic Press.

Lilienfeld, A., Levin, M., and Kessler, I. 1972. Cancer in the United States. Cambridge: Harvard University Press.

Lilienfeld, A.M. and Lilienfeld, D.E. (2ed eds.). 1980. Foundations of Epidemiology. New York: Oxford University Press.

Listort, J.A. 1990. Environmental Health Components for Water Supply, Sanitation and Urban Projects. Washington D.C.: The World Bank Washington, D.C.

Loslier, Luc. 1994. Geographical Information Systems (GIS) from a Health Perspective. In GIS for Health and the Environment. Proceedings of an International Workshop, Colombo, Sri Lanka, D. del Savigny, P. Wijeyaratne, et al.:13-20. Ottawa: International Development Research Center.

Louisiana Department of Environmental Quality. 1987, 1988, 1989, 1990, 1991, 1992, 1993. Louisiana Toxic Release Inventory. Baton Rouge: Louisiana Department of Environmental Quality

Louisiana Department of Environmental Quality, Air Quality Compliance Division. 1992. Annual Toxics Emissions Report. Baton Rouge: Louisiana Department of Environmental Quality.

Louisiana Department of Environmental Quality, Office of the Secretary. 1993. Corporate Response Challenge '92. Baton Rouge: Louisiana Department of Environmental Quality.

Louisiana Department of Environment Quality, Office of Solid and Hazardous Waste (Solid Waste Division). 1993. Final Report Louisiana's Statewide Integrated Solid Waste Management Plan. Baton Rouge: Louisiana Department of Environmental Quality.

Louisiana Department of Environment Quality. 1992. Office of Water Resources. Water Quality Inventory. Baton Rouge: Louisiana Department of Environmental Quality.

Louisiana Department of Environmental Quality, Office of Water Resource, Ground Water Protection Division. 1991. State Profile Document. Baton Rouge: Louisiana Department of Environmental Quality.

Louisiana Department of Labor, Office of Employment Security Research and Statistics Unit. Louisiana Employment and Wages(1980-1994).

Louisiana Department of Public Works. 1971. Ground Water Resources and Requirements for Louisiana: 1970-2020. Baton Rouge: State of Louisiana Department of Public Works.

Louisiana Department of Transportation and Development, Office of Public Works. 1984. The Louisiana Water Resources Study Commissions Report to the 198 Legislature.

Louisiana FactBook. 1993. New Orleans: College of Business Administration, University of New Orleans.

Louisiana State University Medical Center. 1993. River Parishes Cancer Incidence Study: 1983-1987. In Louisiana Toxics Release Inventory, 1992:53-54. Baton Rouge: Department of Environmental Quality.

Lynch, J., Hanis, N., and Bird, M. 1979. An Association of Upper Respiratory Cancer with Exposure to Dimethyl Sulfate. Journal of Occupational Medicine 21:333-341.

MacDonald, E.J. 1976. Demographic Variation in Cancer in Relation to Industrial and Environmental Influence. Environmental Health Perspectives 17:153-166.

Maclur, K.M. and Macmahon, B. 1980. An Epidemiologic Perspective of Environmental Carcinogenesis. Epidemiologic Reviews 2:14-48K.

Mantel, N. 1967. The Detection of Disease Clustering and a Generalized Regression Approach. Cancer Research 27:209-220.

Mary Bird Perkins Cancer Center. 1993. Cancer and Cancer Prevention in South Louisiana. Baton Rouge: Mary Bird Perkins Cancer Center.

Mason, T.J. and Mckay, F.W. 1974. United States Cancer Mortality by County: 1950-1969. Department of Health Education and Welfare Publication No.(NIH) 74-615.

Mason, T.J., Mckay, F.W., Hoover, R., Blot, W.J., and Fraumeni, J.F. 1975. Atlas of Cancer Mortality for U.S. Counties: 1950-1969. Department of Health Education and Welfare Publication No.(NIH) 75-780. Washington, D.C.: U.S. Government Printing Office.

Mason, T.J. Mckay, F.W., Hoover, R., Blot, W.J., and Fraumeni, J.F. 1976. Atlas of Cancer Mortality Among U.S. Nonwhites: 1950-1969. Department of Health Education and Welfare Publication No.(NIH) 76-1204. Washington, D.C.: U.S. Government Printing Office.

Matthews, S. 1990. Epidemiology Using a GIS: the Need for Caution. Computer, Environment and Planning 17(16):213-221.

May, J.M. 1950. Medical Geography : Its Method and Objectives. Geographical Review 40(1):9-41.

Mayer, J.D. 1983. The Role of Spatial Analysis and Geographic Data in the Detection of Disease Causation. Social Science and Medicine 17(16):1213-1221.

McGlashan, N.D and Bond, D.H. 1967. A Method for Using Computer Assistance in Mapping Geographical Synoptic Data. Die Erde 98:292-297.

McGlashan, N.D. 1972. European Male Stomach Cancer in South Africa: A Cartographic Appraisal. In Medical Geography: Techniques and Field Studies, ed. N.D. McGlashan:187-195. London: Methuen and Co Ltd.

McGlashan, N.D. 1983. The Use of Cluster Analysis with Mortality Data. In Geographical Aspects of Health, eds. N.D. MaGlashan and J.R. Blunden, pp. 322-334. London: Academic Press.

McNeil, M. et al. 1981. Environment and Health. Bethesda, Maryland: Congressional Quarterly Inc. National Cancer Institute.

Mead, M.S., Florin, J., and Gesler, W. 1988. Medical Geography. New York: The Guilford Press.

Miller, B.A., Gloeckler Ries, L.A., Hankey, B.F., et al.(eds.). 1993. SEER Cancer Statistics Review 1973-1990. Bethesda: National Cancer Institute.

Minowa, M., Shigematsu, I., et al. 1981. Geographical Distribution of Lung Cancer Mortality and Environmental Factors in Japan. Social Science and Medicine 15D:225-231.

Moore, T.J. 1991. Application of GIS Technology to Air Toxics Risk Assessment: Meeting the Demands of the California Air Toxics 'Hot Spots' Act of 1987. GIS/LIS '91 Proceedings 2:69-714.

Moran, PAP. 1950. Notes on Continuous Stochastic Phenomena. Biometrika 37:17-23.

Moran E.M. 1992. Epidemiological Factors of Cancer in California. Journal of Environmental Pathology, Toxicology, Oncology 11:303-307.

Morgan, R., Foliart, D., and Wong, O. 1985a. Asbestos and Gastrointestinal Cancer: a Review of the Literature. Western Journal Medicine 143:60-65.

Morgan, R. Wong, O. 1985b. A Review of Epidemiological Studies on Artificial Sweeteners and Bladder Cancer. Food Chemical Toxicology 23:529-533.

Mueller, C.W., and Kim, J. 1979. Introduction to Factor Analysis: What it is and How to Do it. London: Sage Publications.

Muir, C.S. and Nectoux, J. 1982. International Patterns of Cancer. In Cancer Epidemiology and Prevention, eds. D. Schottenfeld and F. Farumeni:119-137. Philadelphia: Saunders.

Muir, C., Waterhouse, J., Mack, T., Powell, J., and Whelan, S. (eds.). 1987. Cancer Incidence in Five Continents vol. V. IARC Scientific Publication No.88. Lyon: International Agency for Research on Cancer.

Mulvey, J. and Rothchild, H. 1983. Sugarcane farming - Is There a Link with cancer? In Ecology of Disease, ed. A. Bourne, pp. 267-269. New York: Pergamon.

Murray, M.A. 1962. The Geography Distribution of Cancer Mortality in Wales. Annals of the American Association of Geographers 52:130-149.

National Cancer Institute, Division of Cancer Prevention and Control. 1988. 1987 Annual Cancer Statistics Review. Maryland:NIH Publication No. 88-2789.

Neil. W., et al. 1991. Water Use in Louisiana 1990. Louisiana: Department of Transportation and Development.

New York State Health Department. 1994. GIS Tracks Long Island Breast Cancer. GIS World 7(7):9.

Nickerson, Charles A. and Nickerson, Ingeborg A. 1978. Statistical Analysis for Decision Making, Petrocell Book, Inc., New York, Princeton.

Nicholson, W., Seidman, H., Selikoff, I., et al. 1984. Brain Tumors among Operating Engineers in the Chemical and Petrochemical Industry in Texas and Louisiana. Annals of NY Academic Society 381:172-180.

Normura, A. 1982. Stomach. In Cancer Epidemiology and Prevention, eds. D.Schottenfeld and J.F. Jr. Fraumeni:624-637. Philadelphia: W.B. Saunders.

Norusis, Marija J. and SPSS Inc. 1992. SPSS/PC+ Base System User's Guide Version 5.0/ SPSS/PC+ Advanced Statistics. Chicago: SPSS INC.

Office of Planning and Budget. 1993. Louisiana Benchmarks Supplement Report: Census '90. Baton Rouge: Office of Planning and Budget, Division of Administration.

Office of Planning and Budget. 1994. Louisiana Benchmarks, Economic Edition: a Summary of Economic data for the State of Louisiana. Baton Rouge: Office of Planning and Budget, Division of Administration.

Olsen, G., Hearn, S., Currier, M., Allen, S., and Copk, R. 1989. Mortality Experience of a Cohort of Louisiana Chemical Workers. Journal of Occupational Medicine 31:32-34.

Openshaw, S. 1977. A Geographical Solution to Scale and Aggregation Problems in Region-building, Partitioning and Spatial Modelling. Transactions, Institute of British Geographers, (New Series) 2, 459-472.

Openshaw, S., Charlton, M., Wyner, C., and Craft, A. 1987. A Mark I Geographical Analysis Machine for the Automated Analysis of Point Data Sets. International Journal of Geographic Information Systems 1:335-358.

Openshaw, S., Charlton, M. and Craft, A. 1988. Searching for Leukaemia Clusters Using a Geographical Analysis Machine. Papers of the Regional science Association 64:95-106.

Openshaw, S. 1990. Spatial Analysis and Geographical Information system: A Review of Progress and Possibilities. In Geographical Information Systems for Urban and Regional Planning, eds. M.J. Scholten and J.C.M. Stillwell:153-163. Kluwer, Dordrecht.

Page, H.S. and Asire, A.J. 1985. Cancer Rates and Risks. 3rd edition. National Institutes of Health. NIH Publication No. 85-691.

Page, T., Harris R., and Epstein, S. 1976. Drinking Water and Cancer Mortality in Louisiana. Science 193:55-57.

Phillips, D.R. 1990. Health Care in the Third World. London: Longman.

Pickle, L.W. and Gottlieb, M.S. 1980. Pancreatic Cancer Mortality in Louisiana. American Journal of Public Health 70:256-259.

Pickle, L.W., Mason, T.F., Howard, N., Hoover, R. and Fraumeni, J.F. 1987. Atlas of U.S. Cancer Mortality Among Whites: 1950-1980. Department of Health Education and Welfare Publication No. Washington, D.C.: U.S. Government Printing Office.

Pickle, L.W., Mason, T.F., Howard, N., Hoover, R., and Fraumeni, J.F. 1990. Atlas of U.S. Cancer Mortality Among Nonwhites: 1950-1980. Department of Health Education and Welfare Publication No. Washington, D.C.: U.S. Government Printing Office.

Pollack, E. and Horm, J. 1980. Trends in Cancer Incidence and Mortality in the United States, 1969-76. Journal of National Cancer Institute 64(5):1091-1103.

Potter, J.D. 1996. Nutrition and Colorectal Cancer. Cancer Causes Control 7:157-177.

Public Data Access, Inc. 1988. Mortality and Toxic Wastes along the Mississippi, a Greenpeace report. Washington, D.C.: Greenpeace.

Pyle, G.F. 1968. Some Aspects of Urban Medical Geography. Unpublished Master's thesis, Chicago: University of Chicago, Department of Geography.

Pyle, G.F. 1971. Heart Disease, Cancer, and Stroke in Chicago: a Geographical Analysis with Facilities Plans for 1980. Research Paper No.134. Chicago: University of Chicago, Department of Geography.

Pyle, G.F. and Lauer, B.M. 1975. Comparing Spatial Configurations: Hospital Service Areas and Disease Rates. Economic Geography 51:50-68.

Rainey, J.M., Harwood, A.R., and Crouch, M. 1994. Cancer of the Esophagus in Acadiana. Journal of the Louisiana State Medical Society 146:159-162.

Ranta, J., Karvonen, M. et al. 1996. Detection of Overall Space-Time Clustering in Non-Uniformly Distributed Population. Statistics in Medicine 15:2565-2572.

Ricketts, T.C. et al. 1994. Geographic Methods for Health Services Research. New York: University Press of America.

Riggan W.B, John Van Brugger, John F. Acquavella, Jeff Beaubier, and Mason T.J. 1983. U.S. Cancer Mortality Rates and Trends, 1950-1979, Vol I,II. Washington, D.C.: US Government Printing Office.

Riggan W.B, Mason T.J., et al. 1987. U.S. Cancer Mortality Rates and Trends, 1950-1979, Vol III. Washington, D.C.: US Government Printing Office.

Rothschild, H., Voors, A.W., Weed, S., et al. 1979. Trends in Respiratory System Cancer Mortality in Louisiana: Geographic Distribution in 1950-69 and 1967-76 Compared. American Journal Public Health 69:380-381.

Rothschild, H. and Mulvey, J.J. 1982. An Increase Risk for Lung Cancer Mortality Associated with Sugar Cane Farming. Journal National Cancer Institute 68:755-60.

Rushton, G. and Lolonis, P. 1996. Spatial Analysis of Birth Defect Rates in an Urban Population. Statistics in Medicine 7:717-726.

Schiffman, M.H., Pickle, L.W., and et al. 1988. Case-control Study of Diet and Mesothelioma in Louisiana. Cancer Research 48:2911-2915.

Scholten, H.J., de Lepper Marion, J.C. 1991. The Benefits of the Application of Geographical Information Systems in Public and Environmental Health. World Heath statistics Quartely (Rapport Trimestrie Statistique Sanitaires Mondiales) 44:160-170.

Scott, L.C. 1993. The Chemical Industry in Louisiana: Economic Profile. Baton Rouge: Louisiana Chemical Association.

Scroggins JR., T.G. and Bartley, T.K. 1999. Enhancing Cancer Control: Assessing Cancer Knowledge, Attitudes, and Beliefs in Disadvantaged Communities. Journal of the Louisiana State Medical Society 151(4): 202-208.

Selvin. S., Schulman, J., and Merrill, D.W. 1992. Distance and Risk Measures for Spatial Data: A Study of Childhood Cancers. Social Science and Medicine. 34(7):769-777.

Shafer, S. 1980. Mapping Bone Cancer Death Rates in Pennsylvania Counties. Social Science and Medicine 14D:11-15.

Shannon, G. W. and Pyle, G. 1992. Disease and Medical Care in the United States: A Medical Atlas of the Twentieth Century. New York: Macmillan Pub. Co.

Shear, C., Seale, D. and Gottlieb, M. 1980. Evidence for Space-time Clustering of Lung Cancer Deaths. Archives of Environmental Health 35:335-343.

Snow, J. 1855. On the mode of communication of cholera. London.

Sobel. L.A. 1979. Cancer and the Environment. New York: Facts on File, Inc.

Society for Investigating the Nature and Cure of Cancer. 1806. Edinburgh Medical and Surgical Journal 2:382-389.

Soskolne, C., Zeighami, E., Hanis, N. et al. 1984. Laryngeal Cancer and Occupational Exposure to Sulfuric Acid. American Journal of Epidemiology 120:358-369

Steelman, V.P. 1974. The Cultural Context of Food: A Study of Food Habits and Their Social Significance in Selected Areas of Louisiana, Bulletin No. 679. Baton Rouge: Louisiana State University Center for Agricultural Sciences and Rural Development.

Stimson, R.J. 1983. Research Design and Methodological Problems in the Geography of Health. In Geographical Aspects of Health, ed. N.D. McGlashan and J.R. Blunden. London: Academic Press.

Stone, R.A. 1988. Investigation of Excess Environmental Risk Around Putative Sources: Statistical Problems and a Proposed Test. Statistics in Medicine 7:649-660.

Sutton. P.M. 1965. The Nature of Cancer. New York: Crowell Company.

Tango, T. 1995. A Class of Tests for Detecting General and Focused Clustering of Rare Diseases. Statistics in Medicine 14:2323-2334.

Task Force on Environmental Health. 1984. Environment and Health in Louisiana: The Cancer Problem. Baton Rouge: Louisiana State Planning Office.

Tate, T., Harris, R.H., and Epstein, S.S. 1976. Drinking Water and Cancer Mortality in Louisiana. Science 193:55.

Thomas, T., Stolley, P., Stemhagen, A., et al. 1986. Brain Tumor Mortality Risk among Men with Electrical and Electronics Jobs: a Case-control Study. Journal National Cancer Institute 79:233-238.

Tobacco Institute. 1994. The Tax Burden on Tobacco, Historical Compilation (28). Washington, DC: Tobacco Institute.

Turnbull, B. W., Iwano E. J., Burnett W. S., Howe H. L. and Clark L. C. 1990. Monitoring for Clusters of Disease: Application to Leukemia Incidence in Upstate New York. American Journal of Epidemiology 132:S136-S143.

Twigg, L. 1990. Heath Based Geographical Information Systems: Their Potential Examined in the Light of Existing Data Sources. Social Science and Medicine 30(1):143-155.

U.S. Census Bureau. 1985. Current Population Survey.

U.S. Census Bureau. 1997. Vital Statistics of the United States.

U.S. Department of Commerce, Bureau of the Census. 1983. 1980 Census of Population, Vol 1: Characteristics of the Population, Chapter C: General Social and Economic Characteristics: Louisiana.

U.S. Department of Commerce, Bureau of the Census. 1992. 1990 Census of Population and Housing: Summary Social, Economic, and Housing Characteristics: Louisiana.

U.S. Department of Commerce, Bureau of the Census, Data User Services Division. 1984, 1988, 1994. County and City Data Book. Washington, D.C.

U.S. Department of Commerce, Bureau of the Census, Data User Services Division. 1992. USA Counties. Washington, D.C.

U.S. Department of Commerce, Economics and Statistics Administration, Bureau of the Census. 1995. USA Counties 1994 on CD-ROM. Washington, D.C.

U.S. Department of Health and Human Services. 1989. Reducing the Health Consequences of Smoking: 25 Years of Progress--A Report of the Surgeon General:282-283. Washington, DC: U.S. Government Printing Office.

Verhasselt, Y. 1993. Geography of Health: Some Trends and Perspectives. Social Science and Medicine 30(2):119-123.

Voors, A. et al. 1978. Relationship between Respiratory Cancer and Wetlands Residency in Louisiana. Archives of Environmental Health 33:124-129.

Walter, S.D. 1994. A Simple Test for Spatial Pattern in Regional Health Data. Statistics in Medicine 13:1037-1044.

Waterhouse, J., Muir, C., Shanmugaratnam, K., and Powell, J.(eds.). 1982. Cancer Incidence in Five Continents. vol. IV. IARC Scientific Publication No.42 (1-182) Lyon: International Agency for Research on Cancer.

Weill, H., Hughes, J., and Waggenspack, C. 1979. Influence of Dose and Fiber Type on Respiratory Malignancy Risk in Asbestos Cement Manufacturing. The American Review of Respiratory Disease 120:345-354.

White, R.R. 1972. Probability Maps of Leukaemia Mortalities in England and Wales. In Medical Geography: Techniques and Field Studies, ed. N.D. McGlashan:173-186. London: Methuen & Co. Ltd.

Whittemore, A.S., Friend, H., et al. 1987. A Test to Detect Clusters of Disease. Biometrika 74:631-635.

Wiggins, C.L., Becker, T.M., Key, C.R., and Samet J.M. 1989. Stomach Cancer Among New Mexico's American Indians, Hispanic Whites, and Non-Hispanic Whites. Cancer Research 49:1595-1599.

Wilkins, J.R.(III), Reiches, N.A., and Kruse, C.W. 1979. Organic Chemical Contaminants on Drinking Water and Cancer. American Journal of Epidemiology 110:420-448.

Withers (III), B.F., Ferguson, P.W., and Swift, D.A. 1991. Chemicals, Cancer, and Risk Assessment. Journal of the Louisiana State Medical Society 143:33-40.

Wong, O. and Foliart, D. 1989. Final Report: a Review of Cancer Epidemiology in Louisiana. ENSR Health Sciences.

Wong, O. and Foliart, D. 1993. Epidemiological Factors of Cancer in Louisiana. Journal of Environmental Pathology, Toxicology, and Oncology 12(4):171-18.

World health Organization. 1987-1990. World health Statistics Annuals.

Wyner, E. and Gori, G.B. 1977. Contribution of the Environment of Cancer Incidence: an Epidemiologic Exercise. Journal of National Cancer Institute 58:825-832.

Zeigler, R.G., Mayne, S.T, and Swanson, C.A. 1996. Nutrition and Lung Cancer. Cancer Causes Control 7:157-177.

Zwick. P.D. and Latiner S.S. 1993. Radon Potential Mapping: A Florida Methodology. Proceedings of the Thirteenth Annual ESRI User Conference 3:271-281.

WWW Sites

CDC Wonder. http://wonder.cdc.gov/wonder/data/mortJ.shtml

Mable/Geocorr V2.5-Geographic Correspondence Engine. http://www.census.gov/plue.

SaTScan. http://dcp.nci.nih.gov/BB/SaTScan.html

SpaceSat. http://www.spacesat.com

APPENDIX A: STUDIES ON CANCER AND ENVIRONMENT IN LOUISIANA (SECTION 2.3.3)

Cancer	Cause	Author/Year	Contents
Bladder	Smoking, workers exposed to dye, rubber, or leather are at high risk	Dunham et al.(1968) Blot & Fraumeni (1978) Morgan & Wong (1985a) Gottlieb & Pickle (1981) Nicholson et al.(1984)	Pathology study of urinary bladder cancer Bladder cancer mortality rates by county Artificial sweeteners & bladder cancer Bladder cancer mortality in Louisiana Petrochemical workers
Brain	Nervous system tumors or prenatal X-ray	Thomas et al.(1986)	Occupations (electrical, electronics jobs)
Breast	Increases with age, personal history of breast cancer, late age at menopause or at first live birth, no chef, early age at menarche, effect of radiation	Chen et al.(1992b)	Breast cancer & reproductive system
Colon & rectum	Familial polyposis history, high fat, low fiber	Gottlieb & Carr(1981) Fontham et al.(1992)	Cancer & drinking water Gastrointestinal tract cancers
Esophagus	Heavy consumption of alcohol & tobacco	Rainey et al.(1994	Esophagus in Acadia
Laryngeal	Alcohol, asbestos, nickel, mustard gas, smoking	Lynch et al.(1979) Soskolne et al.(1984)	Alcohol manufacturing plant workers & cancer Occupational exposure to sulfuric acid
Liver	Hepatitis B virus		
Leukemia	Ionizing radiation, certain chemicals such as benzene, host factors such as genetic traits and immune status		
Lung	Smoking, exposure to certain asbestos Cement products/ manufacturing industrial substances (such as arsenic, asbestos, ionizing radiation, and radon)	Weill et al. (1979) Blot et al. (1976,1982) Voor et al. (1978) Gottlieb & Stedman (1979) Gottlieb et al.(1979) Gottlieb (1980) Shear et al. (1980) Herman et al. (1981) Hanis et al. (1982)	Asbestos cement manufacturing Industrial correlations Respiratory cancer & wetlands residency Lung cancer & shipbuilding Lung cancer & death certificate analysis Lung cancer & petrochemical plants Lung cancer & residence near industry Chemical plant workers Refinery & chemical plant workers

(APPENDIX A continued)

Lung		Rothschild et al. (1982)	Lung cancer & sugar farmers
		Gottlieb et al. (1982c)	Lung cancer & residence near industry
		Correa & Johnson (1983)	Cancer & lifestyle
		Correa et al. (1983a)	Passive smoking & lung cancer
		Chen et al. (1984)	Cause of lung cancer deaths in whites
		Correa et al. (1984b)	Occupation & life-style factors (smoking)
		Holmes et al. (1986)	Synthetic rubber plant workers
		Ference et al. (1987)	Chemical plant workers
		Hughes et al. (1987)	Asbestos cement product manufacturing
		Correa et al. (1988)	Diet, nutrition & cancer.
		Fontham et al. (1988a,1993)	Tobacco & cancer
		Fontham et al. (1988b)	Dietary vitamins A, C & lung cancer
		Schiffman et al. (1988)	Asbestos & dietary factors
		Olsen (1989)	Chemical plant workers
		Chen et al. (1992a)	Tobacco-related cancer
		Groves et al. (1996)	Cancer corridor in Louisiana
		Chen et al. (1997)	Highlights of cancer incidence in Louisiana
		Chen et al. (1998)	Cancer incidence in the industrial corridor
Oral	Smoking (cigarette, cigar, pipe, & smokeless tobacco), excess use of alcohol		
Ovary	No chef, late at first pregnancy, late menopause, exposure to asbestos		
Pancreas	Smoking, high fat, chronic pancreatitis, diabetes or cirrhosis	Blot et al. (1978)	Geographic correlates of pancreas cancer
		Pickle & Gottlieb (1980)	Pancreatic cancer & residence near Petrochemical plants
		Hanis et al. (1982)	Refinery & chemical plant workers
		Correa et al. (1984a)	Pancreatic cancer & lifestyle/occupation
		Correa et al. (1988)	Diet, nutrition & cancer

(APPENDIX A continued)

Pancreas		Falk et al. (1988) Fontham et al. (1988a)	Life-style risk factors & cancer Tobacco & cancer
Prostate	Cadmium exposures, fat intake		
Skin	Excessive exposure to ultraviolet radiation, fair complexion, occupational exposure to coal tar, pitch, Creosote, arsenic compounds, or radium		
Stomach	Large amount of pickled, salted or smoked foods, nitrates	Hanis et al. (1982) Correa et al. (1985b) Morgan et al. (1985a) Correa et al. (1988) Fontham et al. (1988a) Fontham et al. (1992)	Refinery & chemical plant workers Dietary determinants of gastric cancer Asbestos & gastrointestinal cancer Diet, nutrition, & cancer Tobacco & cancer Gastrointestinal tract cancers
Cervix & uterus	Multiple sex partners, early age at first intercourse, smoking, certain sexually transmitted diseases		
Uterine & corpus	Obesity, diabetes, high blood pressure		

(Compiled by Bark, 1998)

APPENDIX B: AGENCIES FOR DATA (SECTION 3.2)

1. CANCER

LOUISIANA:
(Louisiana) American Cancer Society
Louisiana Cancer and Lung Trust Fund Board
Louisiana Department of Health and Hospitals
Office of Public Health
    Public Health Statistics
    Louisiana Tumor Registry
Louisiana State University Medical Center
Louisiana State Medical Society
Mary Bird Perkins Cancer Center
Tulane Medical School

UNITED STATES:
American Lung Association
National American Cancer Society
National Cancer Institute
U.S. Department of Health and Human Services
Centers for Disease Control and Prevention

2. ENVIRONMENT

LOUISIANA:
Information: all State Agencies (504)342-6600
Louisiana Department of Agriculture and Forestry
Office of Agricultural and Environmental Science
Division of Pesticides and Environmental Programs
Louisiana Department of Culture, Recreation, and Tourism
Office of Cultural Development
Division of Historic Preservation
Division of Archaeology
Louisiana Department of Economic Development
Louisiana Department of Education
Louisiana Department of Environmental Quality
Office of Air Quality and Radiation Protection
    Air Quality Division
    Radiation Protection Division
Office of Legal Affairs and Enforcement
    Inactive and Abandoned Sites Division
    Legal Division
Office of Solid and Hazardous Waste
    Hazardous Waste Division

(APPENDIX B continued)

Solid Waste Division
Underground Storage Tanks
Louisiana Resource Recovery and Development Authority
Office of Water Resource
Groundwater Protection Division
Water Pollution Control Division
Administrative Hearings Division
Emergency Response Coordinator
Technical Program Support Section
Louisiana Department of Health and Hospitals
Office of Public Health
Environmental Epidemiology Section
Infectious Waste Program
Safe Drinking Water Program
Sewerage Program
Office of Alcohol and Drug Abuse
Louisiana Department of Justice - Attorney General's Office
Environmental Enforcement Section and Citizens' Access Unit
Louisiana Department of Labor (Employment and Training)
Office of Employment Security
Research and Statistics Unit
Louisiana Department of Natural Resources
Coastal Management Division
Office of Coastal Restoration
Office of Conservation
Engineering Division
Injection and Mining Division
Pipeline Division
Louisiana Department of Public Safety & Corrections
Transportation and Environmental Safety Section
Louisiana Department of Revenue and Taxation
Louisiana Department of Transportation and Development
Public Hearing & Environmental Section
Water Resources Section
Louisiana Department of Wildlife and Fisheries
Enforcement Division
Wildlife Regulations Information
Office of Fisheries
Freshwater Fisheries Division
Game Division
Habitat Conservation Division
Public Information and Library
Louisiana Legislature

(APPENDIX B continued)

Senate Switchboard
House of Representatives Switchboard
Public Update Legislative System
Louisiana Litter Control and Recycling Commission
Office Of the Governor
Office of Environmental Affairs and Office of Coastal Activities

UNITED STATES:
U.S. Department of Commerce
National Technical Information Service
U.S. Department of the Army
U.S. Army Corps of Engineers
    Northern & Middle Louisiana Regional Office
    Southern Louisiana Regional Office
U.S. Environmental Protection Agency-Region VI (Dallas, Texas)
Air, Pesticides & Toxics Division
    Air Enforcement Branch
    Air Program Branch
    Pesticides and Toxics Branch
Hazardous Waste Management Division
    RCRA Enforcement Branch
    RCRA Permits Branch
    Superfund Enforcement Branch
    Superfund Programs Branch
    Louisiana Superfund Hazardous Wastes Sites Information
Information Resources Branch
Water Management Division
    Drinking Water Protection Program
    Federal Activities Branch
    Permits Program (NPDES permits)
    Underground Injection Control Program
U.S. Environmental Protection Agency (Washington, D.C.)
Asbestos Information
Air & Radiation Information
Community Right-to-Know Information
Drinking Water Information
Hazardous Waste Information
National Response Center Emergency Hotline
Office of Public Information
Pesticide Information
Wetlands Information
U.S. Department of the Interior
U.S. Fish & Wildlife Service (Slidell)

(APPENDIX B continued)

Wildlife Refuge Section
    Enforcement Section
U.S. Geological Survey
    Biological Resources Division (National Wetlands Research Center)
U.S. Nuclear Regulatory Commission
U.S. Occupational Safety & Health Administration

3. A PARTIAL LISTING OF GROUPS WITH ENVIRONMENTAL CONCERNS

Action Against Waste and to Restore the Environment (AWARE)
Alliance for Affordable Energy
Aluminum Workers Local 275
American Lung Association of Louisiana
Ascension Parish Residents
Atchafalaya Delta Society
Bread for the World
Calcasieu League for Environmental Action Now (CLEAN)
Cankton Cleaners of Land, Air, Water
Cedar Grove Community Group
Church Women United in Louisiana
Citizens Action Committee
Citizens Against Fluoridation
Citizens Against Illegal Dumping (CAID)
Citizens Against the Regional Landfill
Citizen Against Nuclear Trash
Citizens for a Clean Environment (CFACE)
Citizens Organized to Protect Our Parish
Coalition to Restore Coastal Louisiana
Concerned Citizens of Avoyelles (CCA)
Concerned Citizens of Concordia
Concerned Citizens of Plaquemines Parish
Concerned Citizens Committee
Concerned Citizens of Cenla (CCC)
Concerned Citizens of Northeast Louisiana
Concerned Community Citizens of Independence
Delta Greens
Destrehan Neighborhood Alliance
Dialogue Journal
East Iberville AWARE
Fisherman and Concerned Citizens
Grand Isle Stewards of the Environment
Greenpeace
Gulf Coast Tenants Organization

(APPENDIX B continued)

Help Our Environment
Help Our Polluted Environment (HOPE)
Human Ecology Action League (HEAL)
Iberville Parish Landfarm Committee
Industrial Workers of the World (IWW)
Injured Workers Union Local 1
Lowans for a Clean Environment
LaBranche Wetlands Coalition
Lake Pontchartrain Basin Foundation
League of Women Voters State Office
Louisiana Audubon Council (statewide)
Louisiana Consumers League
Louisiana Coalition for Tax Justice
Louisiana Environmental Action Network
Louisiana Nature Conservancy
Louisiana Nature and Science Center
Louisiana Wildlife Federation
Mothers Against Pollution
National Council of Jewish Women
Neighbors Assisting Neighbors (NAN)
North Baton Rouge Environmental Association
Oakville Community Action Group
Oil, Chemical & Atomic Workers Union
Orleans Audubon Society
Parents Against Asbestos
Parents Against Cancer
Pontchartrain Area Recycling
Protecting Environmental and Ecological Resources (PEER)
Restore
Recycle Our Available Resources
Red River Group
Residents Environmental Action League
Save Our Homes & Land
Save Our Lakes & Ducks (SOLD)
Save Our Neighborhoods (SON)
Save Our Selves (SOS)
Save The Trees
Seniors with Power United for Rights (SPUR)
Secure Environment for Everyone
Shrewsbury Neighborhood Organization
Sierra Club Delta Chapter (statewide)
Sierra Club Legal Defense Fund
Slidell Working Against Major Pollution (SWAMP)

(APPENDIX B continued)

South Louisiana Against Pollution (SLAP)
Southern Mutual Help Association, Inc.
slidell Community Opposed to Polluted Environment
U.S. Steelworkers Local B394
St. Bernard Citizens for Environmental Quality
Tulane Environmental Law Clinic
Tulane Green Club
Vermillion Association to Protect
W.A.T.E.R.S.
War Against Waste
Welfare Rights/Client Council of Iberville Parish
Women for a Better Louisiana
Young Leadership Council

(Compiled by Bark, 1998)

APPENDIX C: PROGRAM FOR CANCER MORTALITY RATES (SECTION 3.2.1)

```
/*-------------------------------------------------------------------------*/
/* This program calculates age-adjusted cancer mortality rates    */
/*                    by parish & racesex.                         */
/*-------------------------------------------------------------------------*/

libname s 'c:\sascancer\';
libname a 'a:\';

data uspop;
infile'c:\uspop.prn';
input age uspop;

data lung;set s.lung;
 proc sort;by pr racesex age;

data pop;set s.lpoptran;
 if pr='1' then pr='01';
 if pr='2' then pr='02';
 if pr='3' then pr='03';
 if pr='4' then pr='04';
 if pr='5' then pr='05';
 if pr='6' then pr='06';
 if pr='7' then pr='07';
 if pr='8' then pr='08';
 if pr='9' then pr='09';
 proc sort;by pr racesex age;

data last;merge lung(in=l) lop;by pr racesex age;
if l;
 array d{*} d80-d89;
 array p {*} p80-p89;
 array s {*} s80-s89;
 do i=1 to  10;
   s(i) = d(i)/p(i);
 end;
 drop p80 p81 p82 p83 p84 p85 p86 p87 p88 p89
      d80 d81 d82 d83 d84 d85 d86 d87 d88 d89;
proc sort;by age;

data last1;
merge last uspop;by age;
```

(APPENDIX C continued)

```
  sp80 = s80*uspop*100000;
  sp81 = s81*uspop*100000;
  sp82 = s82*uspop*100000;
  sp83 = s83*uspop*100000;
  sp84 = s84*uspop*100000;
  sp85 = s85*uspop*100000;
  sp86 = s86*uspop*100000;
  sp87 = s87*uspop*100000;
  sp88 = s88*uspop*100000;
  sp89 = s89*uspop*100000;

  keep pr racesex sp80 sp81 sp82 sp83 sp84 sp85 sp86 sp87
     sp88 sp89;

  proc sort;by pr;
   proc summary data=last1;
    class racesex;
    var sp80-sp89;
    by pr;
    output out=l1 sum= ;

data lungout;set l1;
 array ay {*} ay80-ay89;
 array sp {*} sp80-sp89;

 do i= 1 to 10;
  ay(i) = sp(i)/1000000;
 end;

data lungout;set lungout;
 if ay80=. then ay80=0;
 if ay81=. then ay81=0;
 if ay82=. then ay82=0;
 if ay83=. then ay83=0;
 if ay84=. then ay84=0;
 if ay85=. then ay85=0;
 if ay86=. then ay86=0;
 if ay87=. then ay87=0;
 if ay88=. then ay88=0;
 if ay89=. then ay89=0;
 ad=(ay80+ay81+ay82+ay83+ay84+ay85+ay86+ay87+ay88+ay89)/10;
 proc print;
```

(APPENDIX C continued)

```
data a.lungout1;set lungout;
 lungad=ad;
Keep  pr racesex ay80 ay81 ay82 ay83 ay84 ay85 ay86 ay87 ay88 ay89
    lungad;

data a.lungout2;set a.lungout1;
 keep pr racesex lungad;
 proc print;
  var pr racesex lungad;
run;
```

APPENDIX D: PROCEDURE FOR CALCULATING THE DISTANCE OF 2 MILES FROM A SUPERFUND (SECTION 3.2.2)

The processes for calculating the distance (a circle) of 2 miles from a superfund site are as follows.

(1) Digitize the property boundary of a superfund site using Intergraph GIS software.

(2) Line clean the boundary using Intergraph Modular GIS Environment (MGE).

a. Endpoints processor (Tools at the MGE windows -> MGE Base Mapper -> Linework processing -> Endpoints processor).

b. Intersection processor (Tools at the MGE windows -> MGE Base Mapper -> Linework processing -> Intersection processor).

(3) Creating Category and Features for a superfund site.

a. Category Builder (Tools at the MGE Basic Administrator -> Category Builder).

b. Feature Schema Builder (Tools at the MGE Basic Administrator -> Feature Schema Builder).

(4) Processing the features.

a. Running Feature Maker to make the existing linework into a boundary (MGE Base Mapper on Application at the Microstation Command Window -> Tools -> Graphics Processing -> Feature Maker).

b. Map Manager (Tools at the MGE windows -> MGE Base Nucleus -> Map Manager).

c. Running Centroid Placer to put a centroid of each polygon (MGE Base Mapper on Application at the Microstation Command Window -> Tools -> Graphics Processing -> Run Centroid Placer).

(5) Drawing a circle of 2 miles from the centroid of a superfund site (Microstation -> Tools -> Main (standard) -> Place circle -> Draw radius 2 miles (3,218 meter)).

APPENDIX E: PROCEDURE FOR CALCULATING THE AREAS OF WETLANDS OF EACH PARISH (SECTION 3.2.2)

(1) Classified habitat files (for 1978 and 1988) of color infrared aerial photography acquired from Louisiana Department of Natural Resource (LDNR) were imported in raster (.img format) files of Erdas Imagine software.

(2) These files were converted to the Universal Transverse Mercator (UTM) with North American 1927 (Geodetic Datum) coordinate system.

(3) After comparing two classified habitat images with existing wetland maps and reports, areas of wetlands by Corwadian's classification were extracted from the classified habitat image in 1988 and recoded as wetland areas (Image Interpreter -> GIS analysis -> Recode).

(4) The areas of wetlands in each parish were calculated in Erdas Imagine Software

a. Open a raster image file of wetland (.img) (Viewer -> File -> Open -> Raster).
b. Open a vector parish (Louisiana) file of ArcInfo format (Viewer -> File -> Open -> Vector -> Properties -> Polygon) on a raster image file.
c. Create AOI (Area of Interest)s to identify areas of wetlands in each parish polygon (Viewer -> AOI/Tools -> Add to AOI).
d. Subset the AOIs from a raster image file to calculate areas of wetlands in each AOI (Interpreter -> Utilities -> Subset -> Fit to AOI).
e. Calculate areas of wetlands in each parish (Viewer -> Raster -> Attribute Editor -> Add area (Square miles)).

To digitize the areas of wetlands

(1) A rectified habitat raster image file (1988) of Erdas Imagine software was exported as Tagged Imaged File (TIF) format (Erdas Imagine -> Export).

(2) Instead of rectifying the TIF image, its coordinate was adjusted using 'heads-up' tools available in Microstation and Base Imager (Intergraph MGE Base Imager -> Tools -> Edit Header -> Transformation Matrix).

(3) TIF raster image was converted to a vector design file (Intergraph MGE Advanced Imager -> Convert Raster To Vector).

APPENDIX F: TEST OF NORMALITY FOR DATA (SECTION 3.3.1)

Test of Normality for the U.S. Cancer Rates by State (1953-1987)

	Kolmogorov-Smirnov[a]			Shapiro-Wilk		
	Statistic	df	Sig.	Statistic	df	Sig.
ALL5387	.093	49	.200*	.980	49	.702
ALLF5387	.105	49	.200*	.968	49	.371
ALLM5387	.076	49	.200*	.979	49	.693
ALLNW5387	.070	49	.200*	.977	49	.591
ALLW5387	.105	49	.200*	.962	49	.240

*. This is a lower bound of the true significance.

a. Lilliefors Significance Correction

Test of Normality for the U.S. Cancer Rates by County (1953-1987)

	Kolmogorov-Smirnov[a]		
	Statistic	df	Sig.
ALL5387	.320	3067	.000
ALLF5387	.316	3067	.000
ALLM5387	.300	3067	.000
ALLNW5387	.438	3067	.000
ALLW5387	.350	3067	.000

a. Lilliefors Significance Correction

Test of Normality for LA Cancer Rates by Parish (1953-1987)

	Kolmogorov-Smirnov[a]		
	Statistic	df	Sig.
ALL5387	.085	64	.200*
ALLF5387	.077	64	.200*
ALLM5387	.064	64	.200*
ALLNW5387	.065	64	.200*
ALLW5387	.069	64	.200*

*. This is a lower bound of the true significance.

a. Lilliefors Significance Correction

APPENDIX G: PROGRAMS FOR COMPUTING SPATIAL CORRELOGRAM (SECTION 3.3.3)

AUTOCORRELATION

```
/* This program calculates Moran's I and Geary's C for spatial    */
/* autocorrelation. Two files, conn matrix and data matrix, need */
/* to be defined first. */

#include <stdio.h>
#include <math.h>
#define n  65                  /* define number of polygons */
#define t  2               /* define number of time periods */


main ()

{
 int i, j, k;
 char input1[10], input2[10];
 FILE *sf1, *sf2;
 float x [n+1] [t+1], mean [t+1], var [t+1], m4 [t+1];
 float b2 [t+1], sum1 [t+1], sum2 [t+1], m [t+1];
 float c [t+1], varmr [t+1], vargr [t+1], stmn [t+1], stmr [t+1];
 float stgn [t+1], stgr [t+1],  b [n+1] [n+1], sx[n+1],sy[n+1];
 float s0, s1, s2, meanm, varmn, meang, vargn;
 float  xvalue;
 int yvalue;

 printf ("Enter input county conn matrix file name:\n");
 scanf ("%s", input1);
 printf ("Enter input data -x file name:\n");
 scanf ("%s", input2);


 /* READ IN AND PRINT OUT THE CONN MATRIX Bij. CALCULATE S0, S1, S2
*/

 sf1 = fopen (input1, "r");
 for (i=1; i<=n; i++)
        for (j=1; j<=n; j++) {
                fscanf (sf1, "%1d", &yvalue);
                b [i] [j] = 1.0*yvalue;
                }
```

(APPENDIX G continued)

```
for (i=1; i<=n; i++) {
  sx[i] = 0.0;
  sy[i] = 0.0;}
 s0 = 0.0;
 s1 = 0.0;
 s2 = 0.0;

 for (i=1; i<=n; i++)
  for (j=1; j<=n; j++) {
   s0 += b[i][j];
   s1 += pow ((b[i][j] + b[j][i]),2.0);
   sx[i] += b[i][j];          /* sx[i] is row totals */
   sy[j] += b[i][j];}         /* sy[j] is column totals */

s1 = s1 / 2.0;
for (i=1; i<=n; i++)
  s2 += pow ((sx[i] + sy[i]), 2.0);
printf ("\n");
printf ("\n");
printf (" s0=%f\n s1=%f\n s2=%f\n", s0, s1, s2);

printf ("\n");

/* READ IN THE DATA -X MATRIX */

sf2 = fopen (input2, "r");
for (i=1; i<=n; i++) {
        for (j=1; j<=t; j++) {
                fscanf (sf2, "%f", &xvalue);
                x[i][j] = xvalue;
                }
        }
fclose (sf1);
fclose (sf2);

/* CALCULATE THE MEAN, VARIANCE, M4, B2 OF X VALUE */

for (j=1; j<=t; j++) {
        mean [j] = 0;
        for (i=1; i<=n; i++)
                mean [j] += x[i][j]/n;
        }
```

(APPENDIX G continued)

```
for (j=1; j<=t; j++) {
        var [j] = 0;
        m4 [j] = 0;
        for (i=1; i<=n; i++) {
                var[j] += pow ((x[i][j] - mean [j]), 2.0)/n;
                m4[j] += pow ((x[i][j] - mean [j]), 4.0)/n;
                }
        b2 [j] = m4 [j] / (var[j]*var[j]);
        printf("T=%d: mean=%7.2f, var=%8.3f, m4=%12.2f, b2=%7.3f\n",
             j, mean[j], var[j], m4[j], b2[j]);
        }

/* CALCULATE MORAN'S I */

for (j=1; j<=t; j++) {
        sum1[j] = 0;
        for (i=1; i<=n; i++)
                for (k=1; k<=n; k++)
                        sum1[j] += b[i][k]*(x[i][j]-mean[j])*
                                (x[k][j]-mean[j]);
        }

printf("\n");
printf("   Moran's I\n");
printf("Time  I value\n");
for (j=1;j<=t; j++)
   {
      m[j] = sum1[j]/(s0*var[j]);
      printf ("%d    %f\n", j, m[j]);
   }

/* CALCULATE GEARY'S C */

for (j=1; j<=t; j++) {
        sum2 [j] = 0;
        for (i=1; i<=n; i++)
                for (k=1; k<=n; k++)
                        sum2[j] +=b[i][k]*pow((x[i][j]-x[k][j]),2.0);
        }

printf("\n");
printf("   Geary's C\n");
printf("Time  C value\n");
```

(APPENDIX G continued)

```
for (j=1; j<=t; j++) {
        c[j]= (n-1)*sum2[j]/(2*s0*n*var[j]);
        printf ("%d    %f\n", j, c[j]);
        }

 /* CALCULATE STANDARDIZED MORAN'S I */

 meanm = (-1.0)/(n-1);
 varmn = (n*n*s1-n*s2+3*s0*s0)/(s0*s0*(n*n-1))-pow(meanm,2.0);
                              /* normality assumption */
 printf ("\n");
 printf ("Meanm=%f, Varmn=%f\n\n", meanm, varmn);
 for (j=1; j<=t; j++) {
        varmr[j] = ( n*((n*n-3*n+3)*s1-n*s2+3*s0*s0) -
                b2[j]*((n*n-n)*s1-2*n*s2 + 6*s0*s0) ) /
                ( (n-1)*(n-2)*(n-3)*s0*s0 )-pow(meanm, 2.0);
                        /* random assumption */
        stmn [j] = (m[j]-meanm)/sqrt(varmn);
        stmr [j] = (m[j]-meanm)/sqrt(varmr[j]);
        }

 /* CALCULATE STANDARDIZED GEARY'S C */

 meang=1.0;
 vargn = ((2*s1+s2)*(n-1) -4*s0*s0) / (2*(n+1)*s0*s0);   /* normal */
 printf ("\n");
 printf ("Meang=%f, Vargn=%f\n\n", meang, vargn);
 for (j=1; j<=t; j++) {
        vargr [j] = ( (n-1)*s1*(n*n - 3*n + 3 - (n-1)*b2[j]) -
                (n-1)*s2*(n*n + 3*n - 6 - (n*n - n + 2 )*
                b2[j])/4 +s0*s0*(n*n-3-(n-1)*(n-1)*b2[j]) ) /
                ( n*(n-2)*(n-3) *s0*s0);
                        /* random */
        stgn [j] = ( c[j] - meang )/sqrt(vargn);
        stgr [j] = ( c[j] - meang )/sqrt(vargr[j]);
        }

 /* PRINT OUT THE RESULTS */

 printf("\n\n");
 printf("THE STANDARDIZED I AND C VALUES\n");
 printf("Time  I(normal) I(random) varmr     C(normal) C(random) vargr\n");
 for (j=1; j<=t; j++)
```

(APPENDIX G continued)

```
    printf("%d    %-10.4f%-10.4f%-10.4f%-10.4f%-10.4f%-9.4f\n",
          j,stmn[j],stmr[j],varmr[j],stgn[j],stgr[j],vargr[j]);


}
```

(APPENDIX G continued)

CONNECTIVITY

```
/* This program converts the output of the ARC AAT to a connection
matrix, with 1st, 2nd, 3rd, ... orders. */

#include <stdio.h>
#define n  65

/* This program converts the output of the ARC AAT to a connection
matrix, with 1st, 2nd, 3rd, ... orders. */

main ()

{
 char input[12], outp1[12], outp2[12], outp3[12], outp4[12];
 char outp5[12], outp6[12], outp7[12], outp8[12];
 int i, j, k, x, y;
 int w [n+1] [n+1], c2 [n+1][n+1], w2 [n+1][n+1];
 int c3 [n+1][n+1], w3 [n+1][n+1];
 int c4 [n+1][n+1], w4 [n+1][n+1];
 int c5 [n+1][n+1], w5 [n+1][n+1];
 int c6 [n+1][n+1], w6 [n+1][n+1];
 int c7 [n+1][n+1], w7 [n+1][n+1];
 int c8 [n+1][n+1], w8 [n+1][n+1];
 int wb2 [n+1][n+1], wb3 [n+1][n+1];
 int wb4 [n+1][n+1], wb5 [n+1][n+1], wb6 [n+1][n+1];
 int wb7 [n+1][n+1], wb8 [n+1][n+1];
 FILE * sf, *df1, *df2, *df3, *df4, *df5, *df6, *df7, *df8;
 printf ("Enter the input AAT file name:\n");
 scanf ("%s", input);
 for (i=1; i<=n; i++)
  for (j=1; j<=n; j++)
    w [i][j] = 0;
 sf=fopen(input, "r");
 if ( feof(sf) != 1) {
   while ( fscanf (sf, "%d%d", &x,&y)==2 )
     if ( x <= 178 && y <= 178 )
        { w [x][y]=1;   /* w[i][j] is the 1st order conn matrix */
       w [y][x]=1;}
   }

 fclose(sf);
```

(APPENDIX G continued)

```
   for (i=1; i<=n; i++)
  for (k=1; k<=n; k++) {
   c2 [i][k] = 0;
   if (i!=k) {
        for (j=1; j<=n; j++)
         c2 [i][k] += w[i][j] * w [j][k];}
   }              /* c2[i][j] is w[i][j]**2 */

printf("\n");
for (i=1; i<=n; i++) {
  for (j=1; j<=n; j++) {
   if (c2[i][j] != 0 && w[i][j]==0)
        w2 [i][j] = 2;
   else
        w2 [i][j] = w [i][j];
  }
  }              /* w2 [i][j] is the 2nd order conn matrix */


for (i=1; i<=n; i++)
  for (k=1; k<=n; k++) {
   c3 [i][k] = 0;
   if (i!=k) {
        for (j=1; j<=n; j++)
         c3 [i][k] += c2[i][j] * w [j][k];}
   }              /* c3[i][j] is w[i][j]**3 */

printf("\n");
for (i=1; i<=n; i++) {
  for (j=1; j<=n; j++) {
   if (c3[i][j] != 0 && w2[i][j]==0)
        w3 [i][j] = 3;
   else
        w3 [i][j] = w2 [i][j];
  }
  }              /* w3 [i][j] is the 3rd order conn matrix */

for (i=1; i<=n; i++)
  for (k=1; k<=n; k++)  {
   c4 [i][k] = 0;
   if (i!=k) {
        for (j=1; j<=n; j++)
         c4 [i][k] += c3 [i][j] * w [j][k];}
```

(APPENDIX G continued)

```
    }           /* c4 [i][j] is w[i][j]**4 */

 printf("\n");
 for (i=1; i<=n; i++) {
  for (j=1; j<=n; j++) {
   if (c4 [i][j] != 0 && w3 [i][j]==0)
       w4 [i][j] = 4;
   else
       w4 [i][j] = w3 [i][j];
   }
  }           /* w4 [i][j] is the 4th order conn matrix */

 for (i=1; i<=n; i++)
  for (k=1; k<=n; k++)
   {
   c5 [i][k] = 0;
   if (i!=k) {
       for (j=1; j<=n; j++)
        c5 [i][k] += c4 [i][j] * w [j][k];}
   }           /* c5 [i][j] is w[i][j]**5 */

 printf("\n");
 for (i=1; i<=n; i++) {
  for (j=1; j<=n; j++) {
   if (c5 [i][j] != 0 && w4 [i][j]==0)
       w5 [i][j] = 5;
   else
       w5 [i][j] = w4 [i][j];
   }
  }           /* w5 [i][j] is the 5th order conn matrix */



 for (i=1; i<=n; i++)
  for (k=1; k<=n; k++)
   {
   c6 [i][k] = 0;
   if (i!=k) {
       for (j=1; j<=n; j++)
        c6 [i][k] += c5 [i][j] * w [j][k];}
   }           /* c6 [i][j] is w[i][j]**6 */

 printf("\n");
```

(APPENDIX G continued)

```
for (i=1; i<=n; i++) {
 for (j=1; j<=n; j++) {
  if (c6 [i][j] != 0 && w5 [i][j]==0)
      w6 [i][j] = 6;
  else
      w6 [i][j] = w5 [i][j];
  }
 }          /* w6 [i][j] is the 6th order conn matrix  */

for (i=1; i<=n; i++)
 for (k=1; k<=n; k++)
  {
  c7 [i][k] = 0;
  if (i!=k) {
      for (j=1; j<=n; j++)
       c7 [i][k] += c6 [i][j] * w [j][k];}
  }            /* c7 [i][j] is w[i][j]**7 */

printf("\n");
for (i=1; i<=n; i++) {
 for (j=1; j<=n; j++) {
  if (c7 [i][j] != 0 && w6 [i][j]==0)
      w7 [i][j] = 7;
  else
      w7 [i][j] = w6 [i][j];
  }
 }          /* w7 [i][j] is the 7th order conn matrix */


for (i=1; i<=n; i++)
 for (k=1; k<=n; k++)
  {
  c8 [i][k] = 0;
  if (i!=k) {
      for (j=1; j<=n; j++)
       c8 [i][k] += c7 [i][j] * w [j][k];}
  }            /* c8 [i][j] is w[i][j]**8 */

printf("\n");
for (i=1; i<=n; i++) {
 for (j=1; j<=n; j++) {
  if (c8[i][j] != 0 && w7 [i][j]==0)
      w8 [i][j] = 8;
```

(APPENDIX G continued)

```
    else
        w8 [i][j] = w7 [i][j];
    }
  }             /* w8 [i][j] is the 8th order conn matrix */

 for (i=1; i<=n; i++)     /* Assign zero to wbp[i][j]  */
  for (j=1; j<=n; j++)
   {
       wb2 [i][j]=0;
       wb3 [i][j]=0;
       wb4 [i][j]=0;
       wb5 [i][j]=0;
       wb6 [i][j]=0;
       wb7 [i][j]=0;
       wb8 [i][j]=0;
   }          /* wbp[i][j] is the pth (p=1,2,...8) order */
                 /* binary conn matrix which does not contain */
                 /* the the lower order conn information.  */

for (i=1; i<=n; i++)
 for (j=1; j<=n; j++)
  {
      switch ( w8[i][j] )
       {
        case 2:
                wb2 [i][j]=1;
                break;
      case 3:
                wb3 [i][j]=1;
                break;
        case 4:
                wb4 [i][j]=1;
                break;
      case 5:
                wb5 [i][j]=1;
                break;
        case 6:
                wb6 [i][j]=1;
                break;
      case 7:
                wb7 [i][j]=1;
                break;
        case 8:
```

(APPENDIX G continued)

```
                wb8 [i][j]=1;
                break;

        }
    }


printf ("Enter the 1st order binary conn output file name:\n");
scanf ("%s", outp1);
printf ("Enter the 2nd order binary conn output file name:\n");
scanf ("%s", outp2);
printf ("Enter the 3rd order binary conn output file name:\n");
scanf ("%s", outp3);
printf ("Enter the 4th order binary conn output file name:\n");
scanf ("%s", outp4);
printf ("Enter the 5th order binary conn output file name:\n");
scanf ("%s", outp5);
printf ("Enter the 6th order binary conn output file name:\n");
scanf ("%s", outp6);
printf ("Enter the 7th order binary conn output file name:\n");
scanf ("%s", outp7);
printf ("Enter the 8th order binary conn output file name:\n");
scanf ("%s", outp8);

df1=fopen(outp1, "w");
df2=fopen(outp2, "w");
df3=fopen(outp3, "w");
df4=fopen(outp4, "w");
df5=fopen(outp5, "w");
df6=fopen(outp6, "w");
df7=fopen(outp7, "w");
df8=fopen(outp8, "w");

for (i=1; i<=n; i++)
 {
  for (j=1; j<=n; j++)
   {
       fprintf (df1, "%d", w[i][j]);
       fprintf (df2, "%d", wb2[i][j]);
       fprintf (df3, "%d", wb3[i][j]);
       fprintf (df4, "%d", wb4[i][j]);
       fprintf (df5, "%d", wb5[i][j]);
       fprintf (df6, "%d", wb6[i][j]);
```

(APPENDIX G continued)

```
            fprintf (df7, "%d", wb7[i][j]);
            fprintf (df8, "%d", wb8[i][j]);
        }
      fprintf (df1, "\n");
      fprintf (df2, "\n");
      fprintf (df3, "\n");
      fprintf (df4, "\n");
      fprintf (df5, "\n");
      fprintf (df6, "\n");
      fprintf (df7, "\n");
      fprintf (df8, "\n");
    }

  fclose(df1);
  fclose(df2);
  fclose(df3);
  fclose(df4);
  fclose(df5);
  fclose(df6);
  fclose(df7);
  fclose(df8);



}
```

(Source: Algorithms originally developed by N. Lam, 1986
and programmed in C by M. Fan, 1993)

APPENDIX H: CANCER MORTALITY RATES* FOR ALL SITES IN THE U.S.: (1953-1987) (SECTION 4.1)

STATE NAME	Total	Males	Females	Whites	Nonwhites
Alabama	155.16	196.64	124.90	152.11	165.85
Arizona	147.71	181.05	120.97	148.72	130.90
Arkansas	148.84	185.73	118.75	148.04	154.62
California	161.73	196.75	137.30	161.91	158.49
Colorado	138.95	166.10	119.63	138.64	148.84
Connecticut	173.28	216.48	143.75	172.19	197.67
Delaware	180.14	226.34	147.58	172.68	233.26
District of Columbia	205.70	272.94	162.08	173.86	236.19
Florida	157.91	198.41	125.96	153.95	188.81
Georgia	153.59	199.57	122.13	149.11	168.86
Idaho	137.49	162.00	116.92	137.78	115.93
Illinois	171.61	211.34	143.20	167.74	207.16
Indiana	164.45	201.88	137.30	161.72	211.53
Iowa	152.08	185.35	127.77	151.76	185.14
Kansas	144.24	176.33	120.85	142.73	180.20
Kentucky	160.37	195.20	133.31	156.71	210.39
Louisiana	172.96	225.11	134.17	166.60	191.11
Maine	171.66	210.03	144.12	171.79	137.99
Maryland	184.31	234.61	149.76	177.32	221.30
Massachusetts	173.39	217.78	145.12	173.28	176.30
Michigan	171.04	211.46	140.57	167.72	202.73
Minnesota	150.89	179.89	129.10	150.78	158.82
Mississippi	152.14	191.51	121.89	150.45	156.54
Missouri	161.15	199.43	133.23	157.07	207.51
Montana	150.70	178.92	126.68	150.45	160.43
Nebraska	150.74	182.54	126.96	149.80	193.91
Nevada	170.71	205.33	140.75	170.95	165.82
New Hampshire	173.77	215.56	145.17	174.05	89.73
New Jersey	183.49	228.39	152.57	181.32	207.65
New Mexico	139.43	161.79	121.77	141.22	113.40
New York	179.20	219.60	151.26	177.41	194.18
North Carolina	149.25	190.68	119.97	143.39	172.90
North Dakota	143.31	166.50	123.40	143.04	165.52
Ohio	173.00	213.72	143.55	169.29	218.52
Oklahoma	152.21	189.94	123.90	152.40	150.64
Oregon	153.40	185.70	128.32	153.59	144.42
Pennsylvania	173.44	213.45	144.98	170.10	218.46
Rhode Island	180.07	229.43	147.24	179.94	183.98
South Carolina	153.68	199.11	122.77	149.78	165.22

(APPENDIX H continued)

South Dakota	145.26	171.25	124.31	144.77	163.46
Tennessee	154.88	194.13	125.69	149.33	189.20
Texas	151.59	191.78	121.75	148.53	175.84
Utah	122.63	147.31	104.34	122.65	122.16
Vermont	166.63	204.40	140.50	166.81	103.47
Virginia	163.68	207.16	132.69	156.37	196.60
Washington	157.77	191.49	132.15	157.99	151.81
West Virginia	159.71	193.83	132.70	158.48	186.66
Wisconsin	158.17	189.42	134.54	157.50	187.94
Wyoming	139.25	165.96	116.92	139.29	137.06

* Deaths per 100,000, adjusted to the age distribution of the 1970 U.S. Census Population.

(Source: National Technical Information Service, 1992)

APPENDIX I: AGE-ADJUSTED CANCER MORTALITY RATES: SIGNIFICANCE TEST RESULTS, 1953-87 (ALL SITES: WHITES) (SECTION 5.1)

	All Sites: White Males			All Sites: White Females		
South Louisiana	Rate	S.E.	Rate Ratio	Rate	S.E.	Rate Ratio
Acadia	242.14	12.35	1.22**	142.88	8.37	1.06
Allen	200.06	17.18	1.00	120.32	12.34	0.89*
Ascension	214.18	15.85	1.08	119.62	10.48	0.89*
Assumption	221.06	21.95	1.11	101.20	13.06	0.75*
Beauregard	193.05	15.43	0.97	131.52	11.96	0.98
Calcasieu	220.03	8.03	1.10**	127.79	5.26	0.95*
Cameron	169.51	25.22	0.85*	115.43	19.53	0.86
East Baton Rouge	205.24	6.53	1.03	125.88	4.12	0.93*
East Feliciana	141.68	17.05	0.71*	77.95	11.48	0.58*
Evangeline	232.96	15.15	1.17**	139.01	10.52	1.03
Iberia	230.51	13.42	1.16**	136.18	8.86	1.01
Iberville	207.98	18.12	1.04	113.23	12.18	0.84*
Jefferson	234.46	5.85	1.18**	131.63	3.54	0.98
Jefferson Davis	224.91	15.77	1.13**	134.11	10.76	1.00
Lafayette	227.34	10.15	1.14**	128.14	6.21	0.95*
Lafourche	222.11	11.74	1.11**	118.69	7.35	0.88*
Livingston	205.66	13.61	1.03	102.47	8.74	0.76*
Orleans/N`Orlean	253.25	4.30	1.27**	144.07	2.67	1.07**
Plaquemines	239.28	23.92	1.20**	123.77	16.09	0.92
Pointe Coupee	209.14	20.23	1.05	127.11	14.53	0.94
St Bernard	261.85	15.83	1.31**	135.29	9.08	1.00
St Charles	207.32	19.88	1.04	126.36	13.46	0.94
St Helena	194.04	33.82	0.97	90.00	20.01	0.67*
St James	222.67	25.87	1.12	120.97	16.02	0.90
St John Baptist	212.93	23.50	1.07	132.15	16.03	0.98
St Landry	233.53	11.65	1.17**	121.84	7.22	0.90*
St Martin	240.60	17.93	1.21**	136.82	11.98	1.02
St Mary	231.74	14.88	1.16**	129.04	9.61	0.96
St Tammany	215.68	10.36	1.08**	132.95	7.29	0.99
Tangipahoa	208.23	10.85	1.05	126.49	7.45	0.94*
Terrebonne	241.51	12.84	1.21**	124.76	7.95	0.93*
Vermilion	231.30	12.39	1.16**	126.35	8.04	0.94*
Washington	204.27	13.12	1.03	118.70	8.68	0.88*
West Baton Rouge	220.42	28.25	1.11	134.18	19.73	1.00
West Feliciana	182.84	37.51	0.92	112.35	28.11	0.83
LA	219.95	1.58	1.10**	127.85	1.03	1.05
U.S.	199.23	0.16	1.00	134.77	0.12	1.00

See Table 5.1 for footnotes.

APPENDIX J: AGE-ADJUSTED CANCER MORTALITY RATES: SIGNIFICANCE TEST RESULTS, 1953-87 (ALL SITES: NONWHITES) (SECTION 5.1)

	All Sites: Nonwhite Males			All Sites: Nonwhite Females		
South Louisiana	Rate	S.E.	Rate Ratio	Rate	S.E.	Rate Ratio
Acadia	257.54	28.61	1.09	156.91	19.77	1.08
Allen	245.60	36.93	1.04	169.17	28.53	1.17
Ascension	218.29	26.29	0.93	153.66	20.08	1.06
Assumption	262.75	34.42	1.12	181.94	26.72	1.26**
Beauregard	233.42	39.20	0.99	136.46	27.50	0.94
Calcasieu	272.24	17.07	1.16**	151.76	11.75	1.05
Cameron	237.52	19.36	1.01	140.92	88.55	0.97
East Baton Rouge	235.06	9.98	1.00	144.33	6.66	1.00
East Feliciana	133.82	18.21	0.57*	88.15	14.12	0.61*
Evangeline	234.92	33.97	1.00	142.60	24.79	0.98
Iberia	304.02	24.94	1.29**	179.25	17.03	1.24**
Iberville	238.52	21.60	1.01	134.31	14.58	0.93
Jefferson	279.75	16.24	1.19**	161.41	11.27	1.11**
Jefferson Davis	267.57	36.70	1.14	173.68	28.28	1.20**
Lafayette	287.85	22.67	1.22**	167.01	14.70	1.15**
Lafourche	287.58	37.48	1.22**	173.80	26.26	1.20**
Livingston	198.99	41.16	0.85	137.18	33.80	0.95
Orleans/N`Orlean	301.83	6.27	1.28**	177.85	4.04	1.23**
Plaquemines	278.04	40.04	1.18**	164.14	31.39	1.13
Pointe Coupee	218.86	23.62	0.93	141.80	17.68	0.98
St Bernard	277.27	62.74	1.18	162.67	42.52	1.12
St Charles	238.73	33.92	1.01	133.00	22.69	0.92
St Helena	120.29	25.85	0.51*	72.69	19.05	0.50*
St James	250.91	30.25	1.07	163.98	21.60	1.13
St John Baptist	276.25	30.13	1.17**	157.67	20.96	1.09
St Landry	248.12	15.63	1.05	152.52	11.18	1.05
St Martin	267.73	29.36	1.14**	152.17	20.39	1.05
St Mary	266.56	22.53	1.13	185.55	17.22	1.28**
St Tammany	226.31	24.23	0.96	154.91	18.62	1.07
Tangipahoa	215.04	17.76	0.91*	129.54	12.62	0.89*
Terrebonne	288.34	28.15	1.23**	152.01	18.33	1.05
Vermilion	286.22	41.09	1.22**	151.76	26.26	1.05
Washington	217.73	21.16	0.93	122.10	14.38	0.84*
West Baton Rouge	221.62	29.88	0.94	169.87	24.34	1.17**
West Feliciana	179.18	29.94	0.76*	135.01	27.10	0.93
LA	242.58	2.60	1.03**	150.84	1.82	1.04**
U.S.	235.26	0.55	1.00	144.90	0.38	1.00

See Table 5.1 for footnotes.

APPENDIX K: AGE-ADJUSTED CANCER MORTALITY RATES: SIGNIFICANCE TEST RESULTS, 1953-87 (LUNG: WHITES) (SECTION 5.1)

	Lung: White Males			Lung: White Females		
South Louisiana	Rate	S.E.	Rate Ratio	Rate	S.E.	Rate Ratio
Acadia	83.43	7.11	1.47**	20.66	3.19	1.42**
Allen	67.57	9.91	1.19**	12.21	3.95	0.84
Ascension	77.09	9.43	1.35**	15.15	3.78	1.04
Assumption	81.9	13.1	1.44**	8.68	3.93	0.60*
Beauregard	60.58	8.53	1.06	15.92	4.19	1.09
Calcasieu	78.61	4.66	1.38**	17.48	1.95	1.20**
Cameron	68.24	15.56	1.20	24.36	9.03	1.67**
East Baton Rouge	67.39	3.64	1.18**	16.91	1.52	1.16**
East Feliciana	36.72	8.50	0.65*	6.51	3.37	0.45*
Evangeline	82.35	8.82	1.45**	21.27	4.07	1.46**
Iberia	85.84	7.98	1.51**	20.22	3.44	1.39**
Iberville	65.84	10.03	1.16	16.68	4.68	1.14
Jefferson	84.59	3.41	1.49**	20.80	1.41	1.43**
Jefferson Davis	72.75	8.79	1.28**	18.06	3.97	1.24
Lafayette	75.04	5.66	1.32**	17.67	2.33	1.21**
Lafourche	75.02	6.66	1.32**	11.39	2.29	0.78*
Livingston	76.49	8.10	1.34**	14.32	3.30	0.98
Orleans/N`Orlean	79.01	2.34	1.39**	16.19	0.89	1.11**
Plaquemines	73.48	12.38	1.29**	16.45	5.53	1.13
Pointe Coupee	58.08	10.31	1.02	14.70	4.96	1.01
St Bernard	96.30	9.28	1.69**	21.99	3.66	1.51**
St Charles	81.06	12.18	1.42**	15.03	4.59	1.03
St Helena	64.89	18.22	1.14	12.27	7.31	0.84
St James	75.76	14.92	1.33**	9.01	4.45	0.62*
St John Baptist	66.33	13.08	1.17	15.49	5.55	1.06
St Landry	73.50	6.33	1.29**	16.29	2.66	1.12
St Martin	82.62	10.27	1.45**	18.56	4.41	1.27
St Mary	79.83	8.54	1.40**	16.52	3.46	1.13
St Tammany	84.70	6.38	1.49**	23.74	3.09	1.63**
Tangipahoa	67.79	6.09	1.19**	14.41	2.53	0.99
Terrebonne	89.07	7.61	1.57**	12.79	2.55	0.88
Vermilion	67.41	6.56	1.18**	13.94	2.68	0.96
Washington	76.01	7.76	1.34**	15.64	3.16	1.07
West Baton Rouge	74.31	16.13	1.31**	11.84	5.82	0.81
West Feliciana	42.41	17.37	0.75	13.21	9.81	0.91
LA	73.77	0.89	1.30**	16.23	6.37	1.11**
U.S.	56.90	0.09	1.00	14.57	0.04	1.00

Note that for ease of comparisons, the section for Lung: White Males is replicated from Table 5.2. See Table 5.1 for footnotes.

APPENDIX L: AGE-ADJUSTED CANCER MORTALITY RATES: SIGNFICIANCE TEST RESULTS, 1953-87 (LUNG: NONWHITES) (SECTION 5.1)

	Lung: Nonwhite Males			Lung: Nonwhite Females		
South Louisiana	Rate	S.E.	Rate Ratio	Rate	S.E.	Rate Ratio
Acadia	70.90	14.97	1.09	21.61	7.43	1.50
Allen	59.20	18.15	0.91	11.70	7.30	0.81
Ascension	56.92	13.51	0.88	14.13	6.26	0.98
Assumption	79.89	19.10	1.23	12.66	7.20	0.88
Beauregard	70.81	21.80	1.09	11.19	7.81	0.78
Calcasieu	85.41	9.53	1.32**	20.48	4.34	1.42**
Cameron	96.46	80.48	1.49	13.03	25.54	0.91
East Baton Rouge	61.06	5.03	0.94	14.99	2.15	1.04
East Feliciana	32.53	8.96	0.50*	5.80	3.80	0.40*
Evangeline	78.97	19.77	1.22	16.99	9.04	1.18
Iberia	83.91	12.95	1.30**	12.15	4.52	0.84
Iberville	58.20	10.74	0.90	14.46	4.77	1.00
Jefferson	97.63	9.52	1.51**	16.71	3.63	1.16
Jefferson Davis	65.58	18.04	1.01	15.19	8.35	1.06
Lafayette	89.30	12.49	1.38**	21.82	5.30	1.52**
Lafourche	69.78	18.29	1.08	19.31	8.94	1.34
Livingston	70.86	24.71	1.09	18.97	12.52	1.32
Orleans/N'Orlean	89.18	3.34	1.38**	17.71	1.28	1.23**
Plaquemines	79.40	21.22	1.23	24.36	12.56	1.69
Pointe Coupee	58.74	12.27	0.91	12.17	5.16	0.85
St Bernard	88.09	34.15	1.36	20.10	14.90	1.40
St Charles	95.86	21.56	1.48**	17.68	8.43	1.23
St Helena	32.27	13.31	0.50*	6.31	5.56	0.44*
St James	79.22	16.91	1.22	12.24	5.88	0.85
St John Baptist	85.19	16.58	1.32**	19.07	7.23	1.33
St Landry	68.80	8.23	1.06	21.23	4.19	1.48**
St Martin	85.14	16.32	1.31**	9.88	5.21	0.69
St Mary	78.47	12.28	1.21**	21.29	5.94	1.48**
St Tammany	85.09	14.79	1.31**	16.42	6.12	1.14
Tangipahoa	57.44	9.23	0.89	11.12	3.72	0.77
Terrebonne	81.45	14.64	1.26**	21.23	6.95	1.48
Vermilion	69.59	20.15	1.07	14.26	8.10	0.99
Washington	52.76	10.50	0.81*	5.36	3.05	0.37*
West Baton Rouge	76.38	17.76	1.18	14.85	7.32	1.03
West Feliciana	51.88	15.84	0.80	16.99	9.75	1.18
LA	66.94	1.36	1.03**	14.00	0.56	0.97
U.S.	64.76	0.29	1.00	14.39	0.12	1.00

Note that for ease of comparisons, the section for Lung: Nonwhite Males is replicated from Table 5.2. See Table 5.1 for footnotes.

APPENDIX M: AGE-ADJUSTED CANCER MORTALITY RATES: SIGNIFICANCE TEST RESULTS, 1953-87 (STOMACH: WHITES) (SECTION 5.1)

	Stomach: White Males			Stomach: White Females		
South Louisiana	Rate	S.E.	Rate Ratio	Rate	S.E.	Rate Ratio
Acadia	12.64	2.86	1.14	6.62	1.78	1.25
Allen	9.05	3.72	0.82	6.12	2.76	1.15
Ascension	11.45	3.65	1.03	6.64	2.47	1.25
Assumption	9.7	4.57	0.88	3.42	2.38	0.64
Beauregard	10.07	3.50	0.91	4.63	2.22	0.87
Calcasieu	9.58	1.74	0.86	3.47	0.87	0.65*
Cameron	6.97	5.33	0.63	3.56	3.17	0.67
East Baton Rouge	7.26	1.24	0.66*	3.58	0.70	0.67*
East Feliciana	5.95	3.56	0.54*	1.95	1.95	0.37*
Evangeline	11.47	3.42	1.04	3.92	1.72	0.74
Iberia	10.34	3.02	0.93	5.29	1.73	1.00
Iberville	12.87	4.59	1.16	2.65	1.84	0.50*
Jefferson	8.60	1.17	0.78*	4.17	0.64	0.79*
Jefferson Davis	13.58	3.95	1.23	5.53	2.17	1.04
Lafayette	12.92	2.52	1.17	4.43	1.16	0.83
Lafourche	2.55	2.95	1.13	5.77	1.62	1.09
Livingston	6.16	2.51	0.56*	2.33	1.32	0.44*
Orleans/N`Orlean	9.99	0.86	0.90*	5.05	0.48	0.95
Plaquemines	13.33	5.92	1.20	4.01	2.72	0.76
Pointe Coupee	10.34	4.73	0.93	6.28	3.19	1.18
St Bernard	9.99	3.14	0.90	3.22	1.43	0.61*
St Charles	7.58	3.88	0.68	7.87	3.41	1.48
St Helena	7.19	5.80	0.65	3.28	3.74	0.62
St James	9.33	5.09	0.84	6.31	3.60	1.19
St John Baptist	11.05	5.64	1.00	6.64	3.61	1.25
St Landry	14.75	2.95	1.33**	6.66	1.67	1.25
St Martin	16.91	4.69	1.53**	6.38	2.56	1.20
St Mary	10.33	3.12	0.93	3.46	1.60	0.65*
St Tammany	5.95	1.70	0.54*	4.20	1.29	0.79
Tangipahoa	11.14	2.57	1.01	4.54	1.38	0.85
Terrebonne	13.31	3.10	1.20	4.72	1.57	0.89
Vermilion	12.28	2.87	1.11	6.77	1.83	1.27
Washington	7.30	2.48	0.66*	4.69	1.71	0.88
West Baton Rouge	15.77	7.81	1.42	4.75	3.82	0.89
West Feliciana	9.55	8.90	0.86	6.55	6.45	1.23
LA	9.53	0.33	0.86	4.49	0.19	0.85
U.S.	11.08	0.04	1.00	5.31	0.02	1.00

See Table 5.1 for footnotes.

APPENDIX N: AGE-ADJUSTED CANCER MORTALITY RATES: SIGNIFICANCE TEST RESULTS, 1953-87 (STOMACH: NONWHITES) (SECTION 5.1)

	Stomach: Nonwhite Males			Stomach: Nonwhite Females		
South Louisiana	Rate	S.E.	Rate Ratio	Rate	S.E.	Rate Ratio
Acadia	22.83	8.50	1.21	7.24	4.13	0.88
Allen	22.68	11.14	1.21	19.15	9.81	2.34**
Ascension	26.40	9.20	1.40	13.15	5.96	1.61
Assumption	41.07	13.68	2.18**	11.14	6.37	1.36
Beauregard	19.78	11.25	1.05	4.20	4.77	0.51
Calcasieu	29.25	5.64	1.55**	9.85	3.03	1.20
Cameron	17.70	34.70	0.94	0.00	0.00	0.00
East Baton Rouge	20.92	3.00	1.11	8.36	1.62	1.02
East Feliciana	15.92	6.27	0.85	8.53	4.21	1.04
Evangeline	32.20	2.50	1.71* *	11.91	7.22	1.45
Iberia	44.49	9.59	2.36**	17.05	5.25	2.08**
Iberville	29.85	7.58	1.59* *	10.84	4.12	1.32
Jefferson	22.86	4.61	1.21	11.19	3.01	1.37
Jefferson Davis	32.87	12.93	1.75**	7.59	6.13	0.93
Lafayette	29.17	7.09	1.55**	10.18	3.61	1.24
Lafourche	41.89	14.28	2.23**	18.57	8.60	2.27**
Livingston	19.56	12.94	1.04	6.78	7.70	0.83
Orleans/N'Orlean	27.23	1.89	1.45**	11.83	1.05	1.44**
Plaquemines	29.51	12.69	1.57	11.48	8.56	1.40
Pointe Coupee	24.84	7.94	1.32	14.32	5.47	1.75**
St Bernard	54.61	29.97	2.90**	17.40	13.93	2.12
St Charles	26.35	11.35	1.40	15.11	7.54	1.84
St Helena	14.41	9.03	0.77	4.24	4.82	0.52
St James	40.68	12.23	2.16**	15.44	6.51	1.89**
St John Baptist	32.45	10.28	1.72**	14.75	6.34	1.80**
St Landry	29.71	5.38	1.58**	12.79	3.26	1.56**
St Martin	28.58	9.58	1.52**	11.91	5.93	1.45
St Mary	33.09	7.90	1.76**	13.49	4.70	1.65**
St Tammany	15.97	6.43	0.85	12.16	5.11	1.48
Tangipahoa	29.27	6.64	1.56**	11.01	3.79	1.34
Terrebonne	37.73	10.20	1.20**	12.67	5.35	1.55
Vermilion	33.58	14.35	1.78**	7.94	5.91	0.97
Washington	18.28	6.18	0.97	7.05	3.46	0.86
West Baton Rouge	22.49	9.47	1.20	15.17	7.49	1.85
West Feliciana	15.87	8.76	0.84	12.43	8.18	1.52
LA	24.22	0.82	1.29**	10.49	0.48	1.28**
U.S.	18.82	0.16	1.00	8.19	0.09	1.00

Note that for ease of comparisons, the section for Stomach: Nonwhite Males is replicated from Table 5.3. See Table 5.1 for footnotes.

APPENDIX O: AGE-ADJUSTED CANCER MORTALITY RATES: SIGNIFICANCE TEST RESULTS, 1953-87 (COLORECTUM: WHITES) (SECTION 5.1)

	Colorectum: White Males			Colorectum: White Females		
South Louisiana	Rate	S.E.	Rate Ratio	Rate	S.E.	Rate Ratio
Acadia	18.46	3.44	0.72*	18.93	3.02	0.94
Allen	12.95	4.36	0.50*	11.96	3.81	0.59*
Ascension	18.18	4.62	0.71*	16.67	3.91	0.82
Assumption	19.85	6.87	0.77	12.62	4.53	0.62*
Beauregard	17.75	4.72	0.69*	15.79	4.08	0.78*
Calcasieu	18.03	2.35	0.70*	16.79	1.92	0.83*
Cameron	10.87	6.53	0.42*	10.49	5.94	0.52*
East Baton Rouge	21.86	2.19	0.85*	16.67	1.51	0.82*
East Feliciana	14.44	5.64	0.56*	9.90	3.93	0.49*
Evangeline	18.18	4.28	0.71*	16.89	3.63	0.84
Iberia	23.03	4.46	0.89	14.25	2.86	0.70*
Iberville	17.73	5.37	0.69*	13.43	4.17	0.66*
Jefferson	25.51	2.00	0.99	17.21	1.31	0.85*
Jefferson Davis	25.37	5.38	0.98	19.74	4.06	0.98
Lafayette	17.43	2.85	0.68*	13.30	2.01	0.66*
Lafourche	17.92	3.38	0.70*	14.34	2.56	0.71*
Livingston	15.63	3.93	0.61*	12.53	3.08	0.62*
Orleans/N'Orlean	29.41	1.48	1.14**	21.31	0.99	1.05**
Plaquemines	21.09	7.50	0.82	19.04	6.76	0.94
Pointe Coupee	22.35	6.97	0.87	15.02	4.92	0.74*
St Bernard	32.21	5.90	1.25**	20.81	3.67	1.03
St Charles	22.01	6.58	0.85	12.19	4.26	0.60*
St Helena	15.22	10.12	0.59*	8.70	6.06	0.43*
St James	27.90	9.32	1.08	21.74	6.64	1.08
St John Baptist	22.41	7.63	0.87	24.38	6.98	1.21
St Landry	20.09	3.54	0.78*	13.81	2.41	0.68*
St Martin	15.44	4.67	0.60*	14.42	3.89	0.71*
St Mary	19.40	4.39	0.75*	19.94	3.82	0.99
St Tammany	20.04	3.21	0.78*	17.57	2.66	0.87
Tangipahoa	19.43	3.33	0.75*	17.11	2.69	0.85*
Terrebonne	19.82	3.72	0.77*	16.06	2.89	0.79*
Vermilion	17.13	3.36	0.66*	13.84	2.65	0.68*
Washington	11.16	3.09	0.43*	13.65	2.91	0.68*
West Baton Rouge	23.17	9.29	0.90	14.16	6.40	0.70
West Feliciana	29.75	15.80	1.15	5.90	6.71	0.29*
LA	21.23	0.50	0.82	16.91	0.37	0.84
U.S.	25.76	0.06	1.00	20.22	0.05	1.00

See Table 5.1 for footnotes.

APPENDIX P: AGE-ADJUSTED CANCER MORTALITY RATES: SIGNIFICANCE TEST RESULTS, 1953-87 (COLORECTUM: NONWHITES) (SECTION 5.1)

	Colorectum: Nonwhite Males			Colorectum: Nonwhite Females		
South Louisiana	Rate	S.E.	Rate Ratio	Rate	S.E.	Rate Ratio
Acadia	17.59	7.39	0.80	19.61	6.98	1.03
Allen	14.59	9.11	0.67	24.62	10.62	1.30
Ascension	16.12	7.10	0.74	14.57	5.98	0.77
Assumption	8.60	6.39	0.39*	14.65	7.49	0.77
Beauregard	19.43	11.16	0.89	20.93	10.69	1.10
Calcasieu	19.8	4.58	0.91	18.74	4.21	0.99
Cameron	22.48	44.06	1.03	13.03	25.54	0.69
East Baton Rouge	19.97	2.93	0.91	17.09	2.30	0.90
East Feliciana	10.91	5.20	0.50*	5.73	3.58	0.30*
Evangeline	11.54	7.65	0.53*	14.41	8.39	0.76
Iberia	14.2	5.51	0.65*	21.89	5.93	1.16
Iberville	14.53	5.31	0.66*	13.79	4.61	0.73*
Jefferson	21.28	4.58	0.97	17.47	3.72	0.92
Jefferson Davis	18.64	9.82	0.85	28.26	11.61	1.49
Lafayette	19.05	6.03	0.87	14.94	4.44	0.79
Lafourche	17.67	9.33	0.81	24.53	9.94	1.29
Livingston	9.63	9.59	0.44*	12.82	10.48	0.68
Orleans/N`Orlean	27.5	1.92	1.26**	24.16	1.50	1.27**
Plaquemines	24.15	11.95	1.10	12.97	8.19	0.68
Pointe Coupee	15.44	6.21	0.71*	12.35	5.13	0.65*
St Bernard	23.08	18.79	1.06	26.03	17.35	1.37
St Charles	12.78	7.58	0.58*	16.42	8.07	0.87
St Helena	2.57	3.57	0.12*	6.42	5.67	0.34*
St James	21.10	8.68	0.96	25.73	8.60	1.36
St John Baptist	23.36	8.73	1.07	16.62	6.69	0.88
St Landry	14.28	3.79	0.65*	12.07	3.16	0.64*
St Martin	14.43	6.95	0.66*	17.67	7.14	0.93
St Mary	19.60	6.10	0.90	26.57	6.46	1.40**
St Tammany	16.94	6.68	0.77	16.00	5.99	0.84
Tangipahoa	11.04	3.91	0.50*	13.64	4.16	0.72*
Terrebonne	16.42	6.74	0.75	13.45	5.43	0.71
Vermilion	13.53	8.43	0.62	16.15	8.60	0.85
Washington	16.73	5.82	0.76	12.95	4.66	0.68*
West Baton Rouge	12.41	7.04	0.57*	13.50	6.89	0.71
West Feliciana	12.25	8.01	0.56*	20.74	10.27	1.09
LA	18.96	0.73	0.87*	17.77	0.63	0.94*
U.S.	21.87	0.17	1.00	18.95	0.14	1.00

See Table 5.1 for footnotes.

APPENDIX Q: AGE-ADJUSTED CANCER MORTALITY RATES: SIGNIFICANCE TEST RESULTS, 1953-87 (BREAST) (SECTION 5.1)

	Breast: White Females			Breast: Nonwhite Females		
South Louisiana	Rate	S.E.	Rate Ratio	Rate	S.E.	Rate Ratio
Acadia	23.47	3.42	0.88	22.44	7.60	0.93
Allen	20.94	5.24	0.78*	19.29	9.54	0.80
Ascension	21.50	4.48	0.80*	21.19	7.67	0.88
Assumption	15.46	5.23	0.58*	36.11	12.08	1.50
Beauregard	21.61	4.88	0.81*	17.45	9.94	0.72
Calcasieu	23.10	2.22	0.86*	19.12	4.09	0.79*
Cameron	20.62	8.26	0.77	17.74	34.77	0.74
East Baton Rouge	25.15	1.82	0.94	22.04	2.59	0.91
East Feliciana	12.88	4.78	0.48*	11.69	5.29	0.48*
Evangeline	17.14	3.69	0.64*	18.74	8.90	0.78
Iberia	24.30	3.76	0.91	27.96	6.70	1.16
Iberville	17.75	4.84	0.66*	21.19	5.93	0.88
Jefferson	25.90	1.55	0.97	21.28	3.91	0.88
Jefferson Davis	21.67	4.37	0.81*	31.75	11.93	1.32
Lafayette	24.48	2.71	0.91	27.06	5.87	1.12
Lafourche	21.98	3.16	0.82*	20.39	8.84	0.85
Livingston	13.33	3.12	0.50*	19.58	12.32	0.81
Orleans/N`Orlean	27.98	1.19	1.04	30.11	1.65	1.25**
Plaquemines	20.38	6.40	0.76	23.87	11.89	0.99
Pointe Coupee	15.77	5.18	0.59*	17.43	6.32	0.72*
St Bernard	25.06	3.82	0.94	31.32	19.04	1.30
St Charles	22.18	5.52	0.83	17.59	8.02	0.73
St Helena	13.24	7.54	0.49*	14.20	8.44	0.59*
St James	17.02	6.15	0.64*	23.48	8.22	0.97
St John Baptist	20.33	6.26	0.76*	22.09	7.94	0.92
St Landry	18.55	2.83	0.69*	24.46	4.49	1.01
St Martin	21.89	4.81	0.82*	25.75	8.20	1.07
St Mary	22.33	3.97	0.83*	30.32	6.98	1.26
St Tammany	21.11	2.90	0.79*	29.35	8.13	1.22
Tangipahoa	23.20	3.22	0.87*	21.33	5.11	0.88
Terrebonne	21.97	3.30	0.82*	20.94	6.81	0.87
Vermilion	22.70	3.43	0.85*	18.72	9.00	0.78
Washington	20.97	3.66	0.78*	22.50	6.21	0.93
West Baton Rouge	20.91	7.79	0.78	26.67	9.61	1.11
West Feliciana	24.65	2.95	0.92	25.85	12.10	1.07
LA	22.90	0.44	0.85	24.13	0.73	1.00
U.S.	26.79	0.05	1.00	24.13	0.16	1.00

See Table 5.1 for footnotes.

APPENDIX R: AGE-ADJUSTED CANCER MORTALITY RATES: SIGNIFICANCE TEST RESULTS, 1953-87 (PROSTATE) (SECTION 5.1)

	Prostate: White Males			Prostate: Nonwhite Males		
South Louisiana	Rate	S.E.	Rate Ratio	Rate	S.E.	Rate Ratio
Acadia	22.84	4.05	1.11	31.84	10.08	0.89
Allen	21.15	5.72	1.02	43.27	15.53	1.21
Ascension	19.15	5.05	0.93	23.89	8.75	0.67*
Assumption	21.89	7.45	1.06	39.73	13.45	1.11
Beauregard	23.84	5.60	1.15	39.83	15.65	1.12
Calcasieu	19.56	2.64	0.95	27.49	5.74	0.77*
Cameron	15.99	8.48	0.77	26.48	38.73	0.74
East Baton Rouge	17.58	2.14	0.85*	32.73	3.93	0.92
East Feliciana	17.67	6.30	0.86	23.16	7.70	0.65*
Evangeline	20.04	4.63	0.97	24.24	10.99	0.68*
Iberia	16.20	3.74	0.78*	39.14	9.27	1.10
Iberville	19.22	5.84	0.93	39.87	8.71	1.12
Jefferson	17.94	1.81	0.87*	26.73	5.38	0.75*
Jefferson Davis	19.45	4.84	0.94	34.78	13.42	0.98
Lafayette	20.22	3.36	0.98	30.24	7.93	0.85
Lafourche	21.68	4.02	1.05	39.99	14.68	1.12
Livingston	18.90	4.52	0.91	31.62	16.72	0.89
Orleans/N`Orlean	21.23	1.33	1.03	39.61	2.44	1.11**
Plaquemines	23.11	8.57	1.12	20.29	11.09	0.57*
Pointe Coupee	12.64	5.25	0.61*	34.42	9.21	0.97
St Bernard	20.02	5.08	0.97	23.67	19.15	0.66
St Charles	12.01	5.23	0.58*	31.73	12.86	0.89
St Helena	26.55	13.16	1.29	20.99	10.71	0.59*
St James	16.44	7.71	0.80	36.52	11.57	1.02
St John Baptist	18.82	7.53	0.91	42.80	12.36	1.20
St Landry	19.63	3.63	0.95	34.92	5.99	0.98
St Martin	22.32	5.85	1.08	49.87	13.41	1.40*
St Mary	20.31	4.80	0.98	26.86	7.26	0.75*
St Tammany	18.97	3.35	0.92	28.90	8.78	0.81
Tangipahoa	21.66	3.66	1.05	27.84	6.46	0.78*
Terrebonne	22.14	4.37	1.07	41.44	11.16	1.16
Vermilion	26.45	4.39	1.28**	57.02	19.03	1.60**
Washington	24.07	4.84	1.17	43.26	9.45	1.21
West Baton Rouge	10.34	6.64	0.50*	27.45	10.39	0.77
West Feliciana	15.51	12.04	0.75	43.61	15.73	1.22
LA	20.43	0.52	0.99	35.85	1.02	1.01**
U.S.	20.66	0.05	1.00	35.64	0.23	1.00

See Table 5.1 for footnotes.

APPENDIX S: CORRELATIONS FOR INDEPENDENT VARIABLES (SECTIONS 5.4 AND 5.5)

	ALL80S	B80S	CR80S	L80S	P80S	S80S	NWP	PD	PCPI	PBPL
ALL80S	1.000	.244	.650**	.750**	.132	.447**	-.018	.198	.421**	-.197
B80S		1.000	.204	.172	-.142	.131	.085	.232	.419**	-.244
CR80S			1.000	.410**	.137	.334**	.234	.189	.314*	.032
L80S				1.000	-.181	.320**	-.290*	.189	.499**	-.486**
P80S					1.000	.021	.562**	-.109	-.214	.535**
S80S						1.000	.191	.022	.224	-.056
NWP							1.000	-.016	-.209	.658**
PD								1.000	.420**	-.220
PD								1.000	.420**	-.220
PCPI									1.000	-.673**
PBPL										1.000
EDS										
AGR_EM										
MIN_EM										
CON_EM										
MAN_EM										
CHM_EM										
TRA_EM										
HS_EM										
EDS_EM										
TTRI										
TCARC										
NPL										
UR_POP										
D_W_M										
D_W_S										
OZONE										

**Correlation is significant at the 0.01 level (2-tailed), * Correlation is significant at the 0.05 level (2-tailed).
Variables: Normalized Variables – See Chapter 3.2.2.
See Table 5.10 for the variables defined

(APPENDIX S continued)

	EDS	AGR_EM	MIN_EM	CON_EM	MAN_EM	CHM_EM	TRA_EM	HS_EM	EDS_EM	TTRI
ALL80S	.153	-.143	.118	.198	-.026	.203	.361**	-.215	-.110	.185
B80S	.404**	-.398**	.079	.039	.195	.132	.292*	.170	.292*	.132
CR80S	.077	.073	.075	-.088	-.110	.054	.238	-.098	-.014	.171
L80S	.213	-.232	.263*	.434**	.074	.259*	.589**	-.197	-.177	.275*
P80S	-.204	.266*	-.316*	-.394**	-.048	-.142	-.422**	-.030	-.050	-.165
S80S	-.041	-.168	.112	.212	.124	.351**	.274*	-.216	.049	.288*
NWP	-.235	.372**	-.413**	-.338**	-.106	.042	-.397**	.071	.126	-.017
PD	.350**	-.179	-.048	.035	-.079	-.019	.391**	.244	.055	.489**
PCPI	.705**	-.527**	.157	.236	.185	.325**	.653**	.268*	.115	.382**
PBPL	-.691**	.751**	-.220	-.506**	-.352**	-.362**	-.634**	-.117	.034	-.347**
EDS	1.000	-.570**	-.176	.123	.156	.250*	.334**	.288*	.335**	.335**
AGR_EM		1.000	-.106	-.291	-.406**	-.265*	-.444**	-.197	-.007	-.208
MIN_EM			1.000	-.018	-.252*	-.253*	.347**	-.190	-.129	-.192
CON_EM				1.000	.269*	.582**	.425**	-.203	-.111	.350**
MAN_EM					1.000	.540**	.118	-.085	-.082	.414**
CHM_EM						1.000	.229	-.087	-.032	.593**
TRA_EM							1.000	-.067	-.072	.339**
HS_EM								1.000	.075	-.035
EDS_EM									1.000	-.069
TTRI										1.000
TCARC										
H_WS										
S_WS										
NPL										
AGR_CH										
WET										
UR_POP										
D_W_S										
OZONE										

(APPENDIX S continued)

	TCARC	H_WS	S_WS	NPL	AGR_CH	WET	UR_POP	D_W_M	D_W_S	NOZONE
ALL80S	.118	.391**	-.074	.037	.109	.339**	.506**	.447**	.284*	.195
B80S	.136	.400**	.171	.182	-.324**	.123	.429**	.213	.166	.237
CR80S	-.050	.271*	-.102	-.114	.176	.305**	.407**	.427**	.333**	.100
L80S	.126	.340**	-.108	.073	-.041	.547**	.447**	.491**	.343**	.122
P80S	-.091	-.172	-.055	-.120	.225	-.361**	-.042	-.130	-.150	.009
S80S	.209	.237	.024	.063	-.027	.432**	.282*	.336**	.329**	.127
NWP	-.042	.051	.108	-.078	.318**	-.221	-.048	.018	-.019	.118
PD	.127	.561**	-.123	-.004	-.163	.218	.401**	.467**	.337**	.149
PCPI	.333**	.644**	-.060	.211	-.334**	.482**	.694**	.494**	.493**	.326**
PBPL	-.280*	-.359**	.029	-.256*	.615**	-.491**	-.383	-.330**	-.280*	-.203
EDS	.330**	.472**	.079	.257*	-.521**	.141	.581**	.296*	.221	.306*
AGR_EM	-.227	-.321**	-.068	-.160	.784**	-.338**	-.409**	-.274*	-.202	-.179
MIN_EM	-.207	-.083	-.179	-.151	-.008	.247	.043	-.114	.069	-.165
CON_EM	.396**	.288*	.001	.347**	-.172	.469**	-.011	.190	.108	.207
MAN_EM	.363**	.174	.140	.248*	-.361**	.236	.020	.241	.092	.190
CHM_EM	.710**	.558**	.088	.613**	-.108	.331**	.099	.326**	.143	.509**
TRA_EM	.178	.469**	-.096	.113	-.398**	.774**	.471**	.569**	.651**	.087
HS_EM	.008	.157	-.066	-.026	-.158	-.273*	.210	-.041	-.017	.070
EDS_EM	.017	.120	.252*	.015	-.091	-.179	.202	-.096	-.013	.211
TTRI	.575**	.570**	.046	.280*	-.144	.524**	.307*	.681**	.334**	.079
TCARC	1.000	.619**	.073	.632**	-.126	.216	.192	.264*	.172	.247*
H_WS		1.000	-.029	.505**	-.127	.350**	.511**	.550**	.348**	.328**
S_WS			1.000	.097	-.096	-.113	.030	-.046	.028	-.038
NPL				1.000	-.065	.031	-.002	-.050	-.101	.366**
AGR_CH					1.000	-.289*	-.113	-.251	-.215	-.069
WET						1.000	.362**	.669**	.669**	-.061
UR_POP							1.000	.503**	.448**	.229
D_W_M								1.000	.595**	.032
D_W_S									1.000	-.053
OZONE										1.000

APPENDIX T: STEPWISE REGRESSION ANALYSIS OF CANCER MORTALITY RATES (SECTION 5.5)

Stepwise Regression Analysis for All Sites

Source of Variation	Sum of Squares	df	F-ratio	P-value
Regression	6746.740	3	13.621	.000
Residual	9906.421	60		
Total	16653.161	63		

Variables in Equation	R sq Cum	R sq chng	F-value	Beta	T-value
Urban population	.256	.256	21.284	.617	5.944
Health service (employee)	.363	.107	17.380	-.328	-3.219
Educational service (employee)	.405	.042	13.621	-.210	-2.062

Stepwise Regression Analysis for Lung Cancer Mortality Rates (All sites combined)

Source of Variation	Sum of Squares	df	F-ratio	P-value
Regression	1774.525	4	16.423	.000
Residual	1593.791	59		
Total	3368.317	63		

Variables in Equation	R sq Cum	R sq chng	F-value	Beta	T-value
Transportation (employee)	.347	.347	32.944	.315	2.367
Agricultural chemicals	.392	.045	19.641	.440	3.868
Person below poverty	.479	.087	18.384	-.469	-3.473
Mississippi River (drinking)	.527	.048	16.423	.267	2.443

Lung Cancer Mortality Rates: White Males

Source of Variation	Sum of Squares	df	F-ratio	P-value
Regression	2605.061	3	9.267	.000
Residual	5622.137	60		
Total	8227.199	63		

Variables in Equation	R sq Cum	R sq chng	F-value	Beta	T-value
Transportation (employee)	.166	.166	12.350	.586	4.902
Agriculture (employee)	.270	.104	11.269	.464	3.575
Manufacturing (employee)	.317	.047	9.267	.238	2.028

Lung Cancer Mortality Rates: White Females

Source of Variation	Sum of Squares	df	F-ratio	P-value
Regression	1298.235	4	10.797	.000
Residual	1773.569	59		
Total	3071.804	63		

Variables in Equation	R sq Cum	R sq chng	F-value	Beta	T-value
Urban population	.246	.246	20.253	.617	5.565
Mining (employee)	.324	.078	14.638	.298	3.003
Surface water (drinking)	.380	0056	12.249	-.294	-2.630
Construction (employee)	.423	.043	10.797	.209	2.092

Lung Cancer Mortality Rates: Nonwhite Males

Source of Variation	Sum of Squares	df	F-ratio	P-value
Regression	18953.597	3	8.940	.000
Residual	42401.654	60		
Total	61355.251	63		

Variables in Equation	R sq Cum	R sq chng	F-value	Beta	T-value
Person below poverty	.172	.172	12.876	-.542	-3.646
Urban population	.236	.065	9.447	.431	3.272
Education status	.309	.072	8.940	-.423	-2.508

(APPENDIX T continued)

(Nonwhite Females)

Source of Variation	Sum of Squares	df	F-ratio	P-value
Regression	2239.170	2	14.683	.000
Residual	4651.265	61		
Total	6890.435	63		

Variables in Equation	R sq Cum	R sq chng	F-value	Beta	T-value
Urban population	.236	.236	19.198	.489	4.652
Construction (employee)	.325	.089	14.683	.293	2.829

Stepwise Regression Analysis for Breast Cancer Mortality Rates (All sites combined)

Source of Variation	Sum of Squares	df	F-ratio	P-value
Regression	128.226	2	10.783	.000
Residual	362.695	61		
Total	490.921	63		

Variables in Equation	R sq Cum	R sq chng	F-value	Beta	T-value
Urban population	.184	.184	14.007	.398	3.591
Agricultural chemicals	.261	.077	10.783	-.279	-2.520

Breast Cancer Mortality Rates: White Females

Source of Variation	Sum of Squares	df	F-ratio	P-value
Regression	400.837	2	10.912	.000
Residual	1120.399	61		
Total	1521.236	63		

Variables in Equation	R sq Cum	R sq chng	F-value	Beta	T-value
Education status	.213	.213	16.747	.425	3.824
Manufacturing (employee)	.263	.051	10.912	.228	2.052

Breast Cancer Mortality Rates: Nonwhite Females

Source of Variation	Sum of Squares	df	F-ratio	P-value
Regression	1402.869	2	9.233	.000
Residual	4634.260	61		
Total	6037.129	63		

Variables in Equation	R sq Cum	R sq chng	F-value	Beta	T-value
Mining (employee)	.161	.161	11.854	.372	3.299
Agricultural chemicals	.232	.072	9.233	-.270	-2.390

Stepwise Regression Analysis for Colorectal Cancer Mortality Rates (All sites combined)

Source of Variation	Sum of Squares	df	F-ratio	P-value
Regression	284.939	3	9.220	.000
Residual	618.059	60		
Total	902.998	63		

Variables in Equation	R sq Cum	R sq chng	F-value	Beta	T-value
Mississippi River (drinking)	.183	.183	13.856	.376	2.962
Agricultural chemicals	.268	.086	11.177	.299	2.706
Urban population	.316	0.47	9.220	.252	2.038

Colorectal Cancer Mortality Rates: White Males

Source of Variation	Sum of Squares	df	F-ratio	P-value
Regression	105.195	1	7.022	.000
Residual	928.865	62		
Total	1034.060	63		

Variables in Equation	R sq Cum	R sq chng	F-value	Beta	T-value
Mississippi River (drinking)	.102	.102	7.022	.319	2.650

(APPENDIX T continued)

Colorectal Cancer Mortality Rates :White Females

Source of Variation	Sum of Squares	df	F-ratio	P-value
Regression	271.697	4	7.374	.000
Residual	543.432	59		
Total	815.129	63		

Variables in Equation	R sq Cum	R sq chng	F-value	Beta	T-value
Agricultural chemicals	.108	.108	7.518	.506	3.963
Per capita income	.237	.129	9.475	.346	2.552
Construction (employee)	.286	.049	8.014	-.334	-2.734
Wetlands	.333	.047	7.374	.273	2.045

Colorectal Cancer Mortality Rates: Nonwhite Males

Source of Variation	Sum of Squares	df	F-ratio	P-value
Regression	2373.734	2	17.461	.000
Residual	4146.304	61		
Total	6520.037	63		

Variables in Equation	R sq Cum	R sq chng	F-value	Beta	T-value
Urban population	.279	.279	24.036	.736	5.870
Education status	.364	.085	17.461	-.357	-2.850

Colorectal Cancer Mortality Rates: Nonwhite Females

Source of Variation	Sum of Squares	df	F-ratio	P-value
Regression	645.471	2	9.2241	.000
Residual	2130.358	61		
Total	2775.829	63		

Variables in Equation	R sq Cum	R sq chng	F-value	Beta	T-value
Wetlands	.176	.176	13.222	.566	4.296
Total Toxics Releases	.233	.057	9.241	-.279	-2.122

Stepwise Regression Analysis for Prostate Cancer Mortality Rates (All sites combined)

Source of Variation	Sum of Squares	df	F-ratio	P-value
Regression	185.693	2	18.298	.000
Residual	309.530	61		
Total	495.223	63		

Variables in Equation	R sq Cum	R sq chng	F-value	Beta	T-value
Nonwhite population	.316	.316	28.641	.507	4.886
Wetlands	.375	.059	18.298	-.249	-2.399

Prostate Cancer Mortality Rates: White Males

Source of Variation	Sum of Squares	df	F-ratio	P-value
Regression	558.966	2	12.659	.000
Residual	1346.767	61		
Total	1905.733	63		

Variables in Equation	R sq Cum	R sq chng	F-value	Beta	T-value
Person below poverty	.246	.246	20.262	.704	4.725
Education status	.293	.047	12.659	.300	2.014

Prostate Cancer Mortality Rates: Nonwhite Males

Source of Variation	Sum of Squares	df	F-ratio	P-value
Regression	1785.829	2	4.957	.000
Residual	10987.950	61		
Total	12773.779	63		

Variables in Equation	R sq Cum	R sq chng	F-value	Beta	T-value
Transportation (employee)	.065	.065	4.319	-.401	-2.980
Urban population	.140	.075	4.957	.310	2.301

(APPENDIX T continued)

Stepwise Regression Analysis for Stomach Cancer Mortality Rates (All sites combined)

Source of Variation	Sum of Squares	df	F-ratio	P-value
Regression	51.861	2	11.429	.000
Residual	138.405	61		
Total	190.266	63		

Variables in Equation	R sq Cum	R sq chng	F-value	Beta	T-value
Wetlands	.186	.186	14.205	.498	4.450
Nonwhite population	.273	.086	11.429	.301	2.688

Stomach Cancer Mortality Rates: White Males

Source of Variation	Sum of Squares	df	F-ratio	P-value
Regression	54.636	1	6.933	.011
Residual	488.600	62		
Total	543.236	63		

Variables in Equation	R sq Cum	R sq chng	F-value	Beta	T-value
Transportation (employee)	.101	.101	6.933	.317	2.633

Stomach Cancer Mortality Rates: White Females

Source of Variation	Sum of Squares	df	F-ratio	P-value
Regression	27.563	2	6.307	.003
Residual	133.298	61		
Total	160.861	63		

Variables in Equation	R sq Cum	R sq chng	F-value	Beta	T-value
Health service (employee)	.097	.097	6.692	-.303	-2.595
Total Toxics Releases	.171	.074	6.307	.272	2.333

Stomach Cancer Mortality Rates: Nonwhite Males

Source of Variation	Sum of Squares	df	F-ratio	P-value
Regression	3656.643	3	21.199	.003
Residual	3449.773	60		
Total	7106.415	63		

Variables in Equation	R sq Cum	R sq chng	F-value	Beta	T-value
Wetlands	.318	.318	28.900	.354	3.460
Mining (employee)	.456	.138	25.524	.481	4.826
Chemical manufacturing(employee)	.515	.059	21.199	.276	2.700

Stomach Cancer Mortality Rates: Nonwhite Females

Source of Variation	Sum of Squares	df	F-ratio	P-value
Regression	100.191	1	8.457	.005
Residual	734.533	62		
Total	834.724	63		

Variables in Equation	R sq Cum	R sq chng	F-value	Beta	T-value
Wetlands	.120	.120	8.457	.346	2.908

APPENDIX U: SPATIAL CLUSTERS OF LUNG CANCER IN LOUISIANA CENSUS TRACTS, 1988-1993 (SECTION 6.1.2)

Type	Cluster	Census Tract
Spatial analysis (SA) for clusters with high rates Most likely cluster (M)	Cases (c): 3142 Expected (e): 2191.4 Log likelihood ratio (llr): 222.1 P-value (p): 0.001	1)220710054.00, 220710065.00, 220710064.00, 220710050.00, 220710055.00, 220710076.05, 220710045.00, 220710071.00, 220710041.00, 220710063.00, 220710076.04, 220710072.00, 220710070.00, 220710037.01, 220710056.04, 220710037.02, 220710075.02, 220710046.00, 220710075.01, 220710044.01, 220710049.00, 220710056.03, 220710033.05, 220710036.00, 220710044.02, 220710040.00, 220710060.00, 220710124.00, 220710035.00, 220710128.00, 220710033.06, 220710132.00, 220710123.00, 220710069.00, 220510223.02, 220710028.00, 220510224.00, 220710048.00, 220510248.00, 220710034.00, 220710039.00, 220710031.00, 220710033.07, 220710094.00, 220710122.00, 220710056.02, 220710103.00, 220710033.02, 220510223.03, 220710030.00, 220710056.01, 220710130.00, 220710029.00, 220710058.00, 220710112.00, 220710086.00, 220710121.01, 220710131.00, 220710027.00, 220710068.00, 220710059.00, 220710127.00, 220510249.00, 220510226.00, 220710076.03, 220710042.00, 220710033.08, 220510223.01, 220710038.00, 220710026.00, 220510225.00, 220710093.01, 220710080.00, 220710047.00, 220710119.00, 220710023.00, 220710021.00, 220710093.02, 220710102.00, 220710129.00, 220710111.00, 220710085.00, 220710020.00, 220710019.00, 220510222.00, 220710033.01, 220710024.01, 220710018.00, 220710126.00, 220710121.02, 220710067.00, 220710092.00, 220510227.00, 220710079.00, 220710022.00, 220510201.01, 220710133.01, 220510247.00, 220710033.04, 220710057.00, 220710084.00, 220510221.02, 220710100.00, 220710091.00, 220710025.03, 220710015.00, 220710101.00, 220710024.02, 220710109.00, 220710117.00, 220510246.00, 220710012.00, 220510221.01, 220710078.00, 220510228.00, 220710082.00, 220710125.00, 220710001.00, 220710033.03, 220710077.00, 220710090.00, 220510201.02, 220710013.01, 220710025.04, 220710083.00, 220710002.00, 220710099.00, 220710014.01, 220710017.03, 220710025.01, 220710107.00, 220710108.00, 220510220.01, 220710014.02, 220710081.01, 220710116.00, 220710120.00, 220510202.01, 220710088.00, 220510220.02, 220710089.00, 220510245.00, 220710013.04, 220710013.03, 220710013.02, 220510244.00, 220710115.00, 220710003.00, 220710096.00, 220510229.00, 220710097.00, 220710081.02, 220710133.02, 220710016.00, 220710025.02, 220710105.00, 220710017.98, 220710106.00, 220710004.00, 220510202.02, 220710114.00, 220710087.00, 220710135.98, 220510219.00, 220710011.00, 220710017.02, 220510230.01, 220510258.79, 220710104.00, 220710017.06, 220710011.99, 220510243.00, 220710136.98, 220710017.01, 220710104.99, 220710006.01, 220710007.01, 220710006.99, 220510259.00, 220510202.03, 220510231.00, 220510218.02, 220710006.02, 220510230.02, 220710006.05, 220510273.00, 220510257.98, 220710008.00,

(APPENDIX U continued)

M		220510274.00, 220710017.99, 220510260.00, 220710009.04, 220510256.00, 220510266.00, 220710009.03, 220710007.02, 220710009.02, 220510218.01, 220510242.00, 220710007.99, 220710017.20, 220710006.03, 220510230.03, 220510252.01, 220510269.00, 220510272.00, 220510263.00, 220710006.04, 220710009.01, 220510253.00, 220510255.00, 220870303.00, 220510232.00, 220710006.13, 220510267.00, 220510217.00, 220510261.00, 220510270.00, 220510268.00, 220510265.00, 220510203.02, 220710017.24, 220510250.02, 220510233.00, 220510241.00, 220510264.00, 220710006.06, 220510203.03, 220510216.00, 220510254.00, 220510252.02, 220710017.22, 220510239.01, 220510234.00, 220510271.00, 220870305.00, 220510262.00, 220510239.02, 220510240.01, 220510240.02, 220510239.03, 220510215.00, 220510250.03, 220870304.00, 220510239.04, 220510278.03, 220510203.01, 220710006.07, 220510237.00, 220710134.98, 220870306.02, 220870306.01, 220510236.00
Secondary clusters (S)	c:567 e:377.0 llr:42.8 p: 0.001	2)220170216.00, 220170212.00, 220170226.00, 220170215.00, 220170210.00, 220170217.00, 220170227.00, 220170225.00, 220170211.00, 220170218.00, 220170209.00, 220170208.00, 220170228.00, 220170232.00, 220170213.00, 220170202.00, 220170214.00, 220170231.00, 220170219.00, 220170233.00, 220170223.00, 220170230.00, 220170224.00, 220170201.00, 220170207.00, 220170204.00, 220170229.98, 220150102.00, 220150108.01, 220170206.00, 220170221.00, 220170252.98, 220170237.00, 220170220.00, 220150114.98, 220170238.00, 220170234.00, 220150103.00, 220150104.00, 220170252.97, 220150107.02
	c:226 e:128.5 llr:30.5 p: 0.001	3)220190003.00, 220190004.00, 220190015.00, 220190002.00, 220190001.00, 220190006.00, 220190005.00, 220190014.00, 220190025.00, 220190005.99, 220190008.00, 220190009.00
	c:593 e:433.1 llr:27.4 p: 0.001	4)220330018.00, 220330019.00, 220330011.04, 220330011.03, 220330017.00, 220330011.02, 220330010.00, 220330020.00, 220330007.02, 220330027.00, 220330038.01, 220330016.00, 220330006.02, 220330023.00, 220330009.00, 220330007.01, 220330036.01, 220330037.01, 220330006.01, 220330026.01, 220330013.00, 220330015.00, 220330038.02, 220330022.00, 220330005.00, 220330035.05, 220330026.02, 220330025.00, 220330004.00, 220330008.00, 220330035.06, 220330002.00, 220330038.04, 220330014.00, 220330012.00, 220330037.02, 220330021.00, 220330036.03, 220330003.00
	c:177 e:97.0 llr:26.7 p: 0.001	5)220730059.00, 220730057.00, 220730058.00, 220730056.00, 220730014.00, 220730055.00, 220730015.00, 220730016.00, 220730053.02, 220730012.00, 220730013.00, 220730054.00, 220730008.00

(APPENDIX U continued)

S	c:40 e:11.1 llr:22.4 p: 0.001	6)221119604.00
	c:159 e:93.9 llr:18.7 p: 0.001	7)220550011.00, 220550009.00, 220550008.00, 220550001.00, 220550012.00, 220550002.00, 220550006.02
	c:107 e:58.7 llr:16.0 p: 0.001	8)220399505.00, 220399506.00, 220399507.00, 220399504.00
	c:70 e:31.9 llr:12.4 p: 0.004	9)220019608.00, 220019610.00
	c:97 e:56.3 llr:12.2 p: 0.006	10)220790121.00, 220790120.00, 220790122.00, 220790126.00, 220790119.00, 220790110.00
	c:56 e:28.89 llr:9.98 p: 0.042	11)221159502.00, 221159504.00
	p: 0.05	12)220450309.00, 220450308.00, 220450312.00, 220450310.00, 220450307.00
	p: 0.141	13)220119604.00, 220119603.00
	p: 0.487	14)220839705.00
	p: 0.545	15)220039503.00
	p: 0.552	16)220430203.00
	p: 0.581	17)220419501.00, 220419503.00, 220839701.00, 220419502.00
	p: 0.642	18)221139508.00, 221139507.00, 221139506.00

(APPENDIX U continued)

S	p: 0.866	19)221010414.00, 221010415.00, 221010413.00
	p: 0.936	20)220599702.00
	p: 0.960	21)220539805.00, 220539806.00, 2220539807.00, 220539804.0
	p: 0.982	22)220979614.00

See Table 6.3.

VITA

Mihye Bark was born in Taegu, Republic of Korea, on April 3, 1963. She graduated from Hyo-Seung Women's High School, entered Kyungpook National University, majored in geography, and earned her bachelor of science degree in February of 1982. She continued her major at the Graduate School, working as a Graduate Assistant. In 1988, she obtained a master of science in geography.

She entered the doctoral program as a candidate for the degree of Doctor of Philosophy in the Department of Geography and Anthropology at Louisiana State University and will be awarded that degree in May of 2000. While pursuing her Doctor of Philosophy degree, she served as a Graduate Assistant in the Computer Mapping Sciences Laboratory and the Computer-Aided Design and Geographic Information Systems Laboratory at the Department of Geography and Anthropology. Also she worked for Capital Region Planning Commission at Baton Rouge as a GIS planner and Management Info. She is married and lives with her husband and son. She is currently working for the Geographic Information System Unit of Ascension Parish in Louisiana.

DOCTORAL EXAMINATION AND DISSERTATION REPORT

Candidate: Mihye Bark

Major Field: Geography

Title of Dissertation: Cancer Mortality and Environment in Louisiana: A Geographical Inquiry

Approved:

Major Professor and Chairman

Dean of the Graduate School

EXAMINING COMMITTEE:

Date of Examination:

24 November 1999